

PREPARED FOR SUBMISSION TO JHEP

## $\eta$ and $\eta'$ decays into lepton pairs

---

Pere Masjuan,<sup>a</sup> Pablo Sanchez-Puertas<sup>a</sup>

<sup>a</sup>*PRISMA Cluster of Excellence and Institut für Kernphysik, Johannes Gutenberg-Universität,  
Mainz D-55099, Germany*

*E-mail:* [masjuan@kph.uni-mainz.de](mailto:masjuan@kph.uni-mainz.de), [sanchezp@kph.uni-mainz.de](mailto:sanchezp@kph.uni-mainz.de)

**ABSTRACT:** In this work, we calculate the branching ratios for the  $\eta(\eta') \rightarrow \bar{\ell}\ell$  decays, where  $\ell = e, \mu$ . These processes have tiny rates in the Standard Model due to spin flip, loop and electromagnetic suppressions, which could make them sensitive to New Physics effects. In order to provide a reliable input for the Standard Model, we exploit the general analytical properties of the amplitude. For that purpose, we invoke the machinery of Canterbury approximants, which provides a systematic description of the underlying hadronic physics in a data-driven fashion. Given the current experimental discrepancies, we discuss in detail the role of the resonant region and comment on the reliability of  $\chi$ PT calculations. Finally, we discuss the kind of new physics which we think would be relevant to account for them.

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Pseudoscalar decays into lepton pairs: state of the art</b>	<b>2</b>
2.1	Experimental status	2
2.2	Theoretical calculations	3
2.3	A dissection of the $P \rightarrow \bar{\ell}\ell$ process	4
<b>3</b>	<b>Canterbury approximants</b>	<b>6</b>
<b>4</b>	<b>Assessing the systematic error</b>	<b>9</b>
<b>5</b>	<b>Final results</b>	<b>13</b>
<b>6</b>	<b>The low-energy description: the <math>\chi</math>PT approach</b>	<b>16</b>
6.1	The $\pi^0$ -exchange contribution to the $2S$ hyperfine-splitting in the muonic hydrogen	18
6.2	Higher order corrections	18
<b>7</b>	<b>New physics contributions</b>	<b>20</b>
<b>8</b>	<b>Conclusions</b>	<b>25</b>
<b>A</b>	<b>Generalities of Canterbury approximants</b>	<b>26</b>
A.1	Performance of Canterbury approximants with the help of toy models	27

---

## 1 Introduction

Pseudoscalar decays into lepton pairs have been subject of continuous study, starting in the fifties with the first paper by Drell [1] on  $\pi^0 \rightarrow e^+e^-$ . Special interest in these processes lies in the fact that, given the small rates which are predicted within the Standard Model (SM), they could be sensitive to New Physics effects [2–5]. The  $\eta$  and  $\eta'$  cases are of particular interest since they access the  $\mu^+\mu^-$  channel as well, opening the possibility of investigating lepton flavor violation. Given the current puzzles existing in the low-energy precision frontier of particle physics —specifically, the long standing discrepancy among the electron and muon anomalous magnetic moments [6, 7], and the most recent proton radius puzzle coming from the different values obtained from electronic- and muonic-hydrogen experiments [8], together with  $\mathcal{R}_K$  and  $\mathcal{R}_{D^{(*)}}$  [6] anomalies, where  $\mathcal{R}_X = \Gamma(B \rightarrow X\mu^+\mu^-)/\Gamma(B \rightarrow Xe^+e^-)$ —, it is of interest to study whether similar puzzles appear in these processes as well.

$\eta \rightarrow e^+e^-$	$\eta \rightarrow \mu^+\mu^-$	$\eta' \rightarrow e^+e^-$	$\eta' \rightarrow \mu^+\mu^-$	$\pi^0 \rightarrow e^+e^-$
$\leq 2.3 \times 10^{-6}$ [15]	$5.8(8) \times 10^{-6}$ [16]	$\leq 5.6 \times 10^{-9}$ [17, 18]	–	$7.48(38) \times 10^{-8}$ [19]

**Table 1.** Experimental results for  $\text{BR}(\eta, \eta' \rightarrow \bar{\ell}\ell)$  and  $\text{BR}(\pi^0 \rightarrow e^+e^-)$ .

Entering this precision region requires a precise description of the underlying hadronic process fully driving this electromagnetic decay (see Fig. 1) as well as an accurate calculation of the loop integral involved, which itself requires a full-energy hadronic description<sup>1</sup>. Moreover, in the loop integral calculation, it has been common to neglect mass effects [9, 10] which become relevant at a certain precision. In Ref. [11], we showed that a possible approach to avoid the inherent uncertainties from widely-used hadronic models in this calculation is provided by Canterbury approximants, which allow for a data-driven description of the hadronic process in a systematic way; this is because our method exploits the analytic properties of the underlying function. Moreover, we used an exact loop-integral numerical evaluation. In the present work, beyond resuming the main properties of the method [11] to extend it to the  $\eta$  and  $\eta'$  cases, we shall discuss a set of novel features coming from the appearance of intermediate hadronic states absent in the  $\pi^0 \rightarrow e^+e^-$ , and show how our method is able to deal with them. Our results, together with the latest radiative corrections [12–14], would pave the way for the precision low-energy frontier of the Standard Model and New Physics searches in these decays.

The article is organized as follows: In Sec. 2, we report on the state-of-the-art experimental measurements and theoretical predictions together with a general description of the main features of the process under discussion. In Sec. 3, we review the basics of our approach based on Canterbury approximants. In Sec. 4, we discuss the role of intermediate hadronic states and the performance of our approach with the help of a toy model. We present our results in Sec. 5. In Sec. 6, we discuss, in the light of our results, the role of chiral perturbation theory ( $\chi$ PT) when calculating such decays, together with a parameterization to obtain the  $\pi^0$ -exchange contribution to the  $2S$  hyperfine-splitting in the muonic hydrogen; we also derive a general formula to account for the difference between the electronic and muonic channels in the SM beyond the leading order  $\chi$ PT estimate. Finally, with our results at hand, we discuss the implications of experimental results on New Physics searches in Sec. 7. We provide our conclusions and summarize our results in Sec. 8.

## 2 Pseudoscalar decays into lepton pairs: state of the art

### 2.1 Experimental status

Pseudoscalar decays into lepton pairs are considered to be rare since their branching ratios (BR) range from  $10^{-9}$  to  $10^{-6}$ . The state-of-the-art experimental measurements on the  $\text{BR}(\eta, \eta' \rightarrow \bar{\ell}\ell)$  with  $\ell = e, \mu$  are collected in Table 1 where we have also included the  $\pi^0 \rightarrow e^+e^-$  for completeness.

<sup>1</sup>There is in addition a subleading  $Z$ -boson contribution which is necessary to be included to match our precision goal.

The  $\eta \rightarrow e^+e^-$  was measured recently by the HADES collaboration in the context of dark photon searches. HADES is a fixed target experiment where a target of hydrogen or niobium is bombarded with protons and the inclusive  $e^+e^-$  invariant-mass distributions are measured. In 2012, the HADES collaboration obtained with their  $p + p$  data  $\text{BR}(\eta \rightarrow e^+e^-) \leq 5.6 \times 10^{-6}$  [20], while in 2013 they obtained, with their  $p + \text{Nb}$  data,  $\text{BR}(\eta \rightarrow e^+e^-) \leq 2.5 \times 10^{-6}$  [15]. The combined outcome of the two measurements results in the PDG lower limit  $\text{BR}(\eta \rightarrow e^+e^-) \leq 2.3 \times 10^{-6}$  [15].

The  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 5.8(8) \times 10^{-6}$  quoted in the PDG [6] is a combination of the measurements performed at SATURNE II [16] and at Lepton-G [21]. The former was based on the  $pd \rightarrow \eta^3\text{He}$  reaction, measured  $114 \pm 14$  events, and resulted in  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 5.7(7)_{\text{stat.}}(5)_{\text{sys.}} \times 10^{-6}$  [16]. The later was based on the reaction  $\pi^-p \rightarrow \eta n$  with the  $\eta$  reconstructed from the  $\mu^+\mu^-$  invariant mass [21], and measured  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 6.5(2.1) \times 10^{-6}$ .

The bound  $\text{BR}(\eta' \rightarrow e^+e^-) \leq 5.6 \times 10^{-9}$  [17, 18] has been recently established after combining the upper bounds  $\Gamma_{\eta' \rightarrow e^+e^-} < 0.0020$  eV and  $\Gamma_{\eta' \rightarrow e^+e^-} < 0.0024$  eV measured with the SND and the CMD-3 detectors [18] at the VEPP-2000  $e^+e^-$  collider in the  $e^+e^- \rightarrow \eta'$  process. The combination of these two bounds together with the total  $\eta'$  width,  $\Gamma_{\eta'} = 0.198(9)\text{MeV}$  [6], yielded the  $\text{BR}(\eta' \rightarrow e^+e^-) \leq 5.6 \times 10^{-9}$ . Let us note that the PDG still provides the old upper limit  $\text{BR}(\eta' \rightarrow e^+e^-) \leq 2.1 \times 10^{-7}$  measured by the ND Collaboration, Novosibirsk, in 1988 [22].

The process  $\eta' \rightarrow \mu^+\mu^-$  is still to be measured. Even though this process could be measured using the same design as the  $\eta \rightarrow \mu^+\mu^-$  done in SATURNE II, the problem now would be the large background from  $\rho, \omega \rightarrow \mu^+\mu^-$ . A different set-up can be explored with the COMPASS experiment at CERN through the Primakoff effect and the high resolution of the muon pairs in the final state<sup>2</sup>. In addition, it has been recently suggested the possibility to measure this decay at LHCb [23].

Finally, for completeness, we also quote the  $\text{BR}(\pi^0 \rightarrow e^+e^-) = 7.48(38) \times 10^{-8}$  [19] measured by the KTeV Collaboration at FermiLab, which dominates the PDG value.

## 2.2 Theoretical calculations

The first theoretical discussion about  $\eta$  decay into a lepton pair was done by Young in [24]. He used a vector meson dominance (VMD) model for the parameterization of the TFF appearing in the decay. A similar but simpler VMD calculation was performed in [25] by Quigg and Jackson, obtaining<sup>3</sup>  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 4.3 \times 10^{-6}$ . For the  $e^+e^-$  channel, they only reported the unitary bound corresponding to  $\text{BR}(\eta \rightarrow e^+e^-) \geq 1.8 \times 10^{-9}$ . As we will discuss later, the intuition of a unitary bound for  $\eta$  and  $\eta'$  decays no longer holds due to the presence of hadronic intermediate states which reduce the final BR.

Later on, duality ideas were used by Babu and Ma to obtain<sup>4</sup>  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 4.89 \times 10^{-6}$  [26]. These ideas were employed as well in quark-loop model estimations by

<sup>2</sup>We thank Jan Friedrich for discussions along these lines.

<sup>3</sup>The original publication reports on the ratio  $\Gamma_{\mu^+\mu^-}/\Gamma_{\gamma\gamma}$  decay modes. To translate it into the BR here quoted, we used the most recent  $\text{BR}(\eta \rightarrow \gamma\gamma) = 39.41\%$  [6].

<sup>4</sup>Ibid, footnote 3.

Ametller and collaborators [27], obtaining  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 1.13 \times 10^{-5}$  [28]. These may be refined by including a model for the  $q\bar{q}$  bound state wave function [29] yielding  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 4.3 \times 10^{-6}$  and  $\text{BR}(\eta \rightarrow e^+e^-) = 6.3 \times 10^{-9}$ .

The ideas from chiral perturbation theory ( $\chi$ PT) were implemented for the first time in these processes by Savage, Luke, and Wise in [30]. The calculation of the loop process in  $\chi$ PT introduces a counterterm that should be fixed by external information. For that reason, the authors used the experimental  $\text{BR}(\eta \rightarrow \mu^+\mu^-)$  as an input, obtaining  $\text{BR}(\eta \rightarrow e^+e^-) = 5(1) \times 10^{-9}$  as an output. This was revised later on by Gómez Dumm and Pich in [31], finding  $\text{BR}(\eta \rightarrow e^+e^-) = 5.8(2) \times 10^{-9}$ . Using SU(3) symmetry and large- $N_c$  arguments, they also quoted  $\text{BR}(\eta' \rightarrow e^+e^-) = 1.5(1) \times 10^{-10}$  and  $\text{BR}(\eta' \rightarrow \mu^+\mu^-) = 2.1(3) \times 10^{-7}$ .

The development of VMD from an effective field theory point of view together with  $\chi$ PT led to the revision in [32] by Ametller et al., which obtained<sup>5</sup>  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 4.5(^{+3}_{-1}) \times 10^{-6}$ .

A similar study was performed by Silagadze in [33] obtaining  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 5.2(1.2) \times 10^{-6}$  and  $\text{BR}(\eta' \rightarrow \mu^+\mu^-) = 1.4(2) \times 10^{-7}$ . Remarkably, this was the first calculation using an exact result for the loop integral without numerical approximations.

The first implementation of the ideas coming from the large- $N_c$  limit of QCD within the  $\chi$ PT framework and resonant saturation prescription was carried out by the Marseille group in [34] and obtained<sup>6</sup>  $\text{BR}(\eta \rightarrow e^+e^-) = 4.5(2) \times 10^{-9}$  and  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 5.5(8) \times 10^{-6}$ . More recently, the implementation of phenomenological and theoretical constraints in models for the TFFs led the Dubna group in Ref. [35] to the estimates  $\text{BR}(\eta \rightarrow e^+e^-) = 4.60(6) \times 10^{-9}$  and  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 5.11(20) \times 10^{-6}$ . Later on, the implementation of certain numerical corrections discussed in [10] —neglecting errors and using a simple TFF model— yielded in [36] the predictions  $\text{BR}(\eta \rightarrow e^+e^-) = 5.19 \times 10^{-9}$ ,  $\text{BR}(\eta \rightarrow \mu^+\mu^-) = 4.76 \times 10^{-6}$ ,  $\text{BR}(\eta' \rightarrow e^+e^-) = 1.83 \times 10^{-10}$  and  $\text{BR}(\eta' \rightarrow \mu^+\mu^-) = 1.24 \times 10^{-7}$ .

The impact of the numerical corrections above show that approximate calculations do not represent a reliable result for the  $\eta$  and  $\eta'$  cases. Consequently, we will employ an exact numerical evaluation in what follows. For completeness, we also include the  $Z$ -boson contribution in our final results since its size is similar to some of the hadronic uncertainties.

### 2.3 A dissection of the $P \rightarrow \bar{\ell}\ell$ process

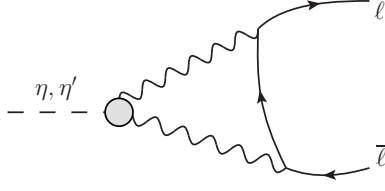
The dominant contribution to the  $P \rightarrow \bar{\ell}\ell$ , illustrated in Fig. 1, is mediated through an intermediate two-photon state, which is a loop process of the momentum  $k$  running through the photons<sup>7</sup>. The main vertex of the process (the blob in Fig. 1), describes the  $P \rightarrow \gamma^*\gamma^*$  transition which is of electromagnetic nature and proceeds through the Adler-Bell-Jackiw anomaly [37, 38]. Since the photons are virtual, the corresponding amplitude reveals the meson structure at all energies in terms of the pseudoscalar-photon transition form factors

---

<sup>5</sup>Ibid, footnote 3.

<sup>6</sup>Ibid, footnote 3.

<sup>7</sup>There is an additional  $Z$  boson contribution to these processes which, as shown in Sec. 7, is subleading but included in our final results.



**Figure 1.** Leading contribution to  $\eta(\eta')$  decays into lepton pairs. The blob represents the TFF.

(TFFs). The TFF, which is a function of the photon virtualities  $q_1^2$  and  $q_2^2$ , is denoted as  $F_{P\gamma^*\gamma^*}(q_1^2, q_2^2)$ .

A description of the TFF at all energies is a formidable task because not only involves different scales but also requires knowledge of the intermediate particles produced by the photons. No complete description of both time-like (TL) and space-like (SL) regions is available so far. Fortunately, due to the kinematics of the process, the required information belongs (mostly) to the SL region, for which certain information is available. For example,  $(F_{P\gamma\gamma}(0,0))^2 \sim \Gamma_{P \rightarrow \gamma\gamma}$ , which are measured quantities and also calculable in the chiral limit; the large photon virtualities are known in turn from perturbative QCD [39]. The interpolation in between these two regimes represents still a challenge in QCD, and demands a model. Since the problem is how to perform the loop calculation in the SL, with scarce set of theoretical input but with great deal of experimental data with respect to the TFF at low and intermediate energies, our attempt in the present work is to perform an interpolation that should satisfy all possible constraints. From the mathematical point of view, this problem is called the *general rational Hermite interpolation problem* and the solution is known to be within the mathematical theory of Padé approximants [40]. Padé approximants (PA) are rational functions  $R_N(x)/Q_M(x)$  with a contact of order  $\mathcal{O}(x^{N+M+1})$  with the function one wants to approximate. That is, their Taylor expansion matches the first  $N + M + 1$  terms from the original function. Only by satisfying these accuracy-through-order conditions can one claim that the analytical properties of the original function are retained. The blob in Fig. 1 represents, however, the TFF of double virtuality and the appropriate approximation to the TFF must be of a bivariate kind. The extension of PA to two variables are called Canterbury approximants [11].

The branching ratio can be expressed, then, in terms of the normalized TFF  $\tilde{F}_{P\gamma^*\gamma^*}(k^2, (q-k)^2)$ , where  $\tilde{F}_{P\gamma^*\gamma^*}(0,0) = 1$ , as

$$\frac{\text{BR}(P \rightarrow \bar{\ell}\ell)}{\text{BR}(P \rightarrow \gamma\gamma)} = 2 \left( \frac{\alpha_{em} m_\ell}{\pi m_P} \right)^2 \beta_\ell |\mathcal{A}(m_P^2)|^2, \quad (2.1)$$

where  $\beta_\ell = (1 - 4m_\ell^2/m_P^2)^{1/2}$  is the outgoing lepton velocity and  $\mathcal{A}(m_P^2)$  is given by the loop integral

$$\mathcal{A}(q^2) = \frac{2i}{\pi^2 q^2} \int d^4k \frac{(q^2 k^2 - (q \cdot k)^2) \tilde{F}_{P\gamma^*\gamma^*}(k^2, (q-k)^2)}{k^2 (q-k)^2 ((p-k)^2 - m_\ell^2)}. \quad (2.2)$$

Anticipating that  $|\mathcal{A}(m_P^2)|$  is typically of  $\mathcal{O}(10)$ , the prefactor in Eq. (2.1) —  $10^{-11}(10^{-7})$

for  $\eta \rightarrow e^+e^-(\mu^+\mu^-)$ , and  $10^{-12}(10^{-7})$  for  $\eta' \rightarrow e^+e^-(\mu^+\mu^-)$ — predicts tiny rates for these processes.

We note that the integral Eq. (2.2) diverges logarithmically for a constant TFF, and requires, at least, a  $k^{-2}$  damping at large energies, which is ensured from the OPE expansion [41]. Though this may suggest a relevant role of the large energies, it turns out that it is the low-energy region (below 1 GeV) of the double virtual TFF which dominates in this calculation [11]. Regretfully, such regime is poorly known and limits the applicability of data-driven approaches.

In order to perform the calculation in Eq. (2.2) —as we have already said— several models for the TFF have been proposed. They started with the advent of VMD ideas [42]. Their theoretical limitations however, arising from the use of a finite set of resonances in a narrow width approximation [43] and lacking the connection with pQCD —a fact circumvented with the use of Padé theory [44, 45]— limited their reliability. To supply this, it has been proposed to use high-energy QCD constraints [26] and, later, well-known phenomenological models constrained from data, which alleviate the model dependency [35]. Moreover, given the lack of double-virtual data on the TFF (which is still an experimental challenge), all the model parameters had to be fixed from high-energy QCD constraints [35] only. Whereas these constraints should be considered, they might not account for the low-energy behavior which, in turn, is the dominant region contributing to this process and may spoil the global description [11].

An alternative idea to circumvent these problems is to use  $\chi$ PT. We only note for the moment that this approach is not well suited though to describe the  $\eta - \eta'$  system and we will comment on its accuracy at the end of Sec. 6.

### 3 Canterbury approximants

In this work, we use the method of Canterbury approximants (CAs) to implement the TFF low-energy behavior together with the QCD constraints in a data-driven approach. Indeed, assuming the relevant intermediate hadronic states to be the  $\rho$  resonance (through  $\pi\pi$  rescattering), and the  $\omega$  and the  $\phi$  narrow-width resonances, all of them with positive imaginary part and belonging therefore to the class of Stieltjes functions, the convergence of CAs would be guaranteed as follows from Refs. [46, 47]. Given a symmetric bivariate function  $F_{P\gamma^*\gamma^*}(Q_1^2, Q_2^2) = F_{P\gamma^*\gamma^*}(Q_2^2, Q_1^2)$  with known Taylor expansion

$$F_{P\gamma^*\gamma^*}(Q_1^2, Q_2^2) = F_{P\gamma^*\gamma^*}(0, 0) \left( 1 - \frac{b_P}{m_P^2}(Q_1^2 + Q_2^2) + \frac{c_P}{m_P^4}(Q_1^4 + Q_2^4) + \frac{a_{P;1,1}}{m_P^4}Q_1^2Q_2^2 + \dots \right), \quad (3.1)$$

CAs [11, 46, 48] are rational functions<sup>8</sup> of bivariate polynomials of degree  $N$  and  $M$ , respectively, which coefficients are defined as to match the low-energy expansion of the original  $F_{P\gamma^*\gamma^*}(Q_1^2, Q_2^2)$  function, i.e.,

$$C_M^N(Q_1^2, Q_2^2) \equiv \frac{R_N(Q_1^2, Q_2^2)}{Q_M(Q_1^2, Q_2^2)}, \quad (3.2)$$

---

<sup>8</sup>We refer the interested reader to Appendix A and to Ref. [11], where CAs were introduced.

fulfilling the conditions specified by Eqs. (A.3, A.4). Only from this definition is our rational function guaranteed to converge to the original function provided it fulfills certain analytical properties (for instance, if it is a meromorphic [40, 44, 45, 49–52] or a Stieltjes function [40, 47, 49, 50, 53, 54]). This construction allows to describe the TFF with the correct low-energy implementation, which is known to play the main role in these processes, a fact often overlooked (see the discussion in [11]). Moreover, it has the ability to implement at the same time the high energy QCD-behavior, which is of relevance for the high-energy tail of the integration.

As discussed in Ref. [11], the simplest element of the  $C_M^N(Q_1^2, Q_2^2)$  sequence reads [11]:

$$C_1^0(Q_1^2, Q_2^2) = \frac{a_0}{1 + \frac{b_P}{m_P^2}(Q_1^2 + Q_2^2) + (\frac{2b_P^2 - a_{P;1,1}}{m_P^4})Q_1^2 Q_2^2}, \quad (3.3)$$

where  $a_0 = F_{P\gamma\gamma}(0, 0)$  is related to the  $\Gamma_{P\gamma\gamma}$ ,

$$|F_{P\gamma\gamma}(0, 0)|^2 = \frac{64\pi}{(4\pi\alpha)^2} \frac{\Gamma(P \rightarrow \gamma\gamma)}{m_P^3}, \quad (3.4)$$

with  $m_P$  the pseudoscalar mass.  $F_{\eta\gamma\gamma}(0, 0) = 0.274(5)\text{GeV}^{-1}$  and  $F_{\eta'\gamma\gamma}(0, 0) = 0.344(6)\text{GeV}^{-1}$  using the  $\Gamma_{P\gamma\gamma}$  from the PDG [6]. The parameter  $b_P$  is the slope of the single-virtual TFF [51, 55, 56] and ensures the appropriate low-energy behavior up to  $\mathcal{O}(Q^4)$  corrections. The most precise determinations for the slope of the  $\eta$  and  $\eta'$  TFFs have been obtained from a data-driven procedure and read  $b_\eta = 0.576(12)$  [55] and  $b_{\eta'} = 1.31(4)$  [56], respectively. In addition, for one virtuality  $C_1^0(Q^2, 0)$  satisfies the Brodsky-Lepage (BL) [39] high-energy  $Q^2$ -behavior by construction. For the  $\pi^0$ , the BL asymptotic limit reads  $\lim_{Q^2 \rightarrow \infty} F_{P\gamma^*\gamma^*}(Q^2, 0) = 2F_\pi Q^{-2}$  with  $F_\pi = 92.21 \text{ MeV}$  [6] the pion decay constant. For the  $\eta$  and  $\eta'$ , a similar formula exists which involves both the  $\eta - \eta'$  mixing parameters and the running effects of the singlet axial current [55, 56]. In our case, when we state that we fulfill the correct BL  $Q^2$ -behavior, we refer to the fact that our approximant satisfies  $\lim_{Q^2 \rightarrow \infty} Q^2 F_{P\gamma^*\gamma^*}(Q^2, 0) = \mathcal{C} Q^{-2}$ , but without fixing the precise  $\mathcal{C}$  coefficient (i.e.  $\mathcal{C} = 2F_\pi$  for the  $\pi^0$  and its counterpart for the  $\eta$  and  $\eta'$ ). Not fixing  $\mathcal{C}$  is extremely important in our method, given the relevance of the low energies when using the simplest  $C_1^0(Q_1^2, Q_2^2)$  approximant to perform the numerical evaluation of Eq. (2.2). Matching  $b_P/m_P^2$  to  $\mathcal{C}$  instead of to the TFF's slope will spoil the TFF low-energy expansion dominant in our decays— see discussion in Appendix A and Table 8.

Not only that, but the resulting description can be checked to reproduce the low-energy data on the single-virtual TFF to an excellent precision and the —less relevant— intermediate and high-energy data up to 10% accuracy, see Figs. 6 and 7 in this respect as well.

Finally,  $a_{P;1,1}$  is the double virtual slope of the TFF. Since there is no experimental data for the double virtual TFF so far, the method of fitting the experimental data with a PA sequence considered in Refs. [51, 55–58] cannot be used to obtain the double-virtual parameters. For this reason, we suggested in Ref. [11] to choose a generous range for this  $a_{P;1,1}$  parameter that should cover the well-known theoretical constraints at low and high



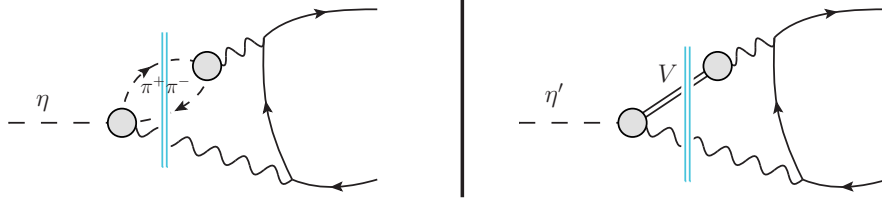
energies, including the yet unknown real value. It turns out that, at low energies—the most important region in our calculation— $\chi$ PT calculations favor the so called factorization approach [59], namely, that  $\tilde{F}_{P\gamma^*\gamma^*}(Q_1^2, Q_2^2) = \tilde{F}_{P\gamma^*\gamma}(Q_1^2, 0) \times \tilde{F}_{P\gamma\gamma^*}(0, Q_2^2)$ , implying that  $a_{P;1,1} = b_P^2$  in Eq. (3.3)<sup>9</sup>. On the other hand, it is known that at high energies the operator product expansion (OPE) dictates that  $F_{P\gamma^*\gamma^*}(Q^2, Q^2) \sim Q^{-2}$  [26, 35], implying  $a_{P;1,1} = 2b_P^2$  in Eq. (3.3)<sup>10</sup>. Deviations from factorization should translate into  $a_{P;1,1} > b_P^2$ . Then, from now on and due to the ignorance on the precise value of  $a_{P;1,1}$ , we will assume for the rest of our work  $b_P^2 \leq a_{P;1,1} \leq 2b_P^2$ . In addition, we note that from Eq. (3.3),  $a_{P;1,1} \leq 2b_P^2$  is required for avoiding poles in the SL region.

To scrutinize a bit more the interplay of these two regimes, let us look briefly at the  $\pi^0 \rightarrow e^+e^-$  decay, for which is possible to employ a reliable approximated version for Eq. (2.2) [11] thanks to the smallness of the  $\pi^0$  mass. This allows to split the integral ( $I$ ) in Eq. (2.2) in two parts as  $I = \int_0^\Lambda \text{Fact} + \int_\Lambda^\infty \text{OPE}$ , in which the low-energy part below the scale  $\Lambda$  is integrated using the factorization approach (i.e.,  $a_{P;1,1} = b_P^2$ ), whereas the high-energy part above  $\Lambda$  is integrated using the OPE ( $a_{P;1,1} = 2b_P^2$ ). We obtain  $\text{BR}(\pi^0 \rightarrow e^+e^-) = \{6.15, 6.21, 6.26, 6.29\} \times 10^{-8}$  for  $\Lambda = \{0, 1, 2, \infty\}$ . This shows that, if factorization represents a good approximation at low energies, the freedom on choosing when the OPE should contribute implies a non-negligible systematic error at the precision we are aiming—of the order of the percent level—and vice-versa. It is reasonable to assume then, that such deviation should be covered by choosing a conservative range  $a_{P;1,1} \in (b_P^2, 2b_P^2)$ . In addition, we remind that the mathematical theory of CAs does not dictate what  $a_{P;1,1}$  to take and both choices are equally correct as they only correspond to different reconstructions. We illustrate these discussions, e.g. the relevant range for the  $a_{P;1,1}$  parameter and the relevance of using the low-energy parameters, in the Table 8 of Appendix A with the help of two well motivated toy models. Again, we emphasize that the reconstructed element,  $C_1^0(Q_1^2, Q_2^2)$ , fulfill the different  $Q^2$ -behaviors which are implied by QCD for  $a_{P;1,1} = 2b_P^2$ , but not the corresponding coefficients—these could be implemented in higher elements, which unfortunately involve at the moment too many unknown parameters of double virtuality. Still, from our studies in [11] together with Appendix A and the reasons presented above, these effects are seen to be tiny and the quoted value already provides a reliable result.

The last ingredient in our approach is the estimation of a systematic error induced by the truncation of the CA sequence, i.e., the fact that even though the CA sequence converges to the TFF, at a given and finite CA order, the difference between the function and its approximant is not zero. The inclusion of a systematic error marks a difference with respect to previous studies. This is the goal of the next chapter.

<sup>9</sup>Actually, the recent results from dispersive analysis [60] seem to confirm the factorization behavior employed in [11] for the  $\eta$  case. However, as we will show later, this approach is only safe at very low energies and the region from  $1 - 2 \text{ GeV}^2$  is already dominated by the pQCD behavior.

<sup>10</sup>Again, here we only fulfill the  $Q^2$ -behavior, i.e., that  $\lim_{Q^2 \rightarrow \infty} F_{P\gamma^*\gamma^*}(Q^2, Q^2) \propto Q^{-2}$ , but we do not impose the particular coefficient, which turns out to be  $1/3$  of the BL behavior [26, 35].



**Figure 2.** Intermediate hadronic states invalidating the unitary bound for the  $\eta(\eta')$ , left(right).

#### 4 Assessing the systematic error

The  $\eta$  and  $\eta'$  masses are large enough to yield intermediate hadronic states in the  $P \rightarrow \bar{\ell}\ell$  processes as sketched in Fig. 2, which implies an additional imaginary part beyond that of the  $\gamma\gamma$  contribution via Cutkosky rules. As we have already said, this novel feature—not present for the  $\pi^0$ —diminishes the imaginary part with respect to the  $\gamma\gamma$  intermediate state invalidating the unitary bound [1],  $|\mathcal{A}(m_P^2)|^2 \geq \text{Im}(\mathcal{A}_{\gamma\gamma}(m_P^2))^2$ , which tacitly assumes the absence of intermediate states beyond the  $\gamma\gamma$  one. This effect can be understood from the loop integral as well. After Wick-rotation, the TFF integration domain lies in the  $-m_P^2 \leq Q^2 \leq \infty$  region, which for the  $\eta$  and  $\eta'$  involves threshold production ( $m_\eta > 2m_\pi$ ) and resonances ( $m_{\eta'} > m_{\rho,\omega} > 2m_\pi$ ).

This effect has never been considered before when calculating the rare decays and must be taken into account as well when evaluating the systematic error. As we will show below, the uncertainty from the resonance part of Fig. 2 becomes the dominant source of error. To quantitatively study this effect, we take a toy model for the TFF that includes both a two-pion production threshold and a vector resonance. The model is conceived in such a way that the time-like region contains all the required features of the physical TFF up to the  $\eta'$  mass. The first ingredient in our toy model is factorization, which as explained before seems a reasonable choice at low energies. The second ingredient is the use of vector meson dominance ideas [61] allowing to express the single-virtual TFF as

$$\tilde{F}_{P\gamma^*\gamma}(s) = c_{P\rho}G_\rho(s) + c_{P\omega}G_\omega(s) + c_{P\phi}G_\phi(s), \quad (4.1)$$

where  $G_V(s)$  are the different resonance contributions weighted by the dimensionless couplings  $c_{PV}$  which are obtained from a quark-model,  $c_{\eta(\eta')\rho} = 9/8(9/14)$ ,  $c_{\eta(\eta')\omega} = 1/8(1/14)$ ,  $c_{\eta(\eta')\phi} = -2/8(4/14)$  [62], and  $G_V(0) = 1$ . In order to incorporate the  $\pi\pi$  intermediate branch cut from Fig. 2, fulfilling unitarity and analyticity, we take for the  $\rho$  contribution  $G_\rho(s)$  a model based on Refs. [63, 64]

$$G_\rho(s) = \frac{M_\rho^2}{M_\rho^2 - s + \frac{sM_\rho^2}{96\pi^2 F_\pi^2} \left( \ln\left(\frac{m_\pi^2}{\mu^2}\right) + \frac{8m_\pi^2}{s} - \frac{5}{3} - \sigma(s)^3 \ln\left(\frac{\sigma(s)-1}{\sigma(s)+1}\right) \right)} \quad (4.2)$$

with  $\sigma(s) = \sqrt{1 - 4m_\pi^2/s}$ , and the parameters  $M_\rho = 0.815$  GeV,  $F_\pi = 0.115$  GeV,  $\mu = 0.775$  GeV, and  $m_\pi = 0.139$  GeV, chosen to reproduce the pole position  $s_\rho = (M - i\Gamma/2)^2$  with

$M = 0.764$  GeV and  $\Gamma = 0.144$  GeV from [65], while for the (narrow-width)  $\omega, \phi$  resonances, we take<sup>11</sup>

$$G_{\omega, \phi} = \frac{M_{\omega, \phi}^2 + M_{\omega, \phi} \Gamma_{\omega, \phi} (s_{th}/M_{\omega, \phi}^2)^{3/2}}{M_{\omega, \phi}^2 - s + M_{\omega, \phi} \Gamma_{\omega, \phi} ((s_{th} - s)/M_{\omega, \phi}^2)^{3/2}}, \quad (4.3)$$

which parameters are fixed from PDG masses and widths [6]. This choice makes our model very similar to the dispersive approach formulated in [62].

To evaluate now the branching ratio of the decay, we have to calculate the diagram in Fig. 1. The blob there stands for the TFF, which includes among others the contributions from Fig. 2 and is accounted for by the model in Eq. (4.1). Employing the Cauchy integral representation for the (factorized) TFF,

$$F_{P\gamma^*\gamma^*}(q_1^2, q_2^2) = \frac{1}{\pi^2} \int_{s_{th}}^{\infty} dM_1^2 \int_{s_{th}}^{\infty} dM_2^2 \frac{\text{Im}[F_{P\gamma^*\gamma}(M_1^2)]}{q_1^2 - M_1^2 - i\epsilon} \frac{\text{Im}[F_{P\gamma^*\gamma}(M_2^2)]}{q_2^2 - M_2^2 - i\epsilon}, \quad (4.4)$$

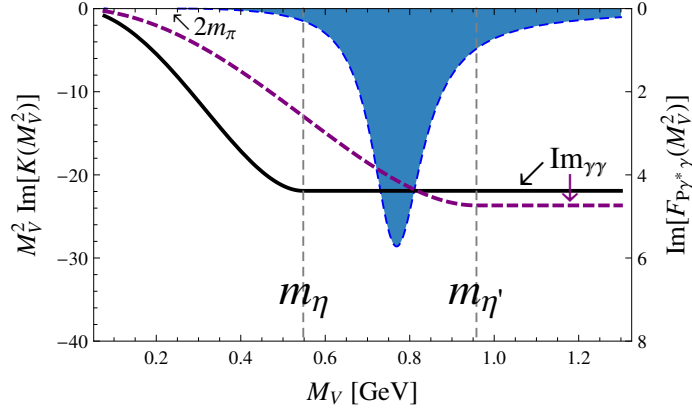
and changing the integration order in the integral (2.2), allows to express the loop amplitude as

$$\begin{aligned} \mathcal{A}(q^2) &= \frac{1}{\pi^2} \int_{s_{th}}^{\infty} dM_1^2 \int_{s_{th}}^{\infty} dM_2^2 \text{Im}[\tilde{F}_{P\gamma^*\gamma}(M_1^2)] \text{Im}[\tilde{F}_{P\gamma^*\gamma}(M_2^2)] \\ &\quad \times \left( \frac{2i}{\pi^2 q^2} \int d^4k \frac{(q^2 k^2 - (qk)^2)}{k^2 (q-k)^2 ((p-k)^2 - m_\ell^2)} \frac{1}{k^2 - M_1^2} \frac{1}{(q-k)^2 - M_2^2} \right) \\ &\equiv \frac{2}{\pi^2} \int_{s_{th}}^{\infty} dM_1^2 \int_{s_{th}}^{M_1^2} dM_2^2 \text{Im}[\tilde{F}_{P\gamma^*\gamma}(M_1^2)] \text{Im}[\tilde{F}_{P\gamma^*\gamma}(M_2^2)] \times K(M_1^2, M_2^2). \end{aligned} \quad (4.5)$$

This procedure results on an easy evaluation of the loop amplitude denoted as  $K(M_1^2, M_2^2)$  through standard one-loop techniques [66] or a numerical evaluation using `LoopTools` [67].

Now, the threshold effects are clear and easier to handle. To illustrate them, we plot the imaginary part of the integrand in (4.5) in terms of  $\text{Im}[\tilde{F}_{P\gamma^*\gamma}(M_V^2)]$  and  $\text{Im}[K(M_V^2)]$  —which contains both  $\gamma\gamma$  and vector contributions— when dispersing only one virtuality in (4.5) for simplicity (i.e., we neglect the  $q^2$  dependence on the second virtuality). The resulting plot is shown in Fig. 3 as a solid-black (dashed-purple) line for the  $\eta(\eta')$  in terms of the dispersive variable  $M_V$  once the  $\int d^4k$  integration has been performed to give  $K(M_V^2)$  in the last line of Eq. (4.5). These lines have to be convoluted with  $\text{Im}[\tilde{F}_{P\gamma^*\gamma}(M_V^2)]$ . In Fig. 3, this is represented by the bluish area. For clarity, we only plot there  $\text{Im}[G_\rho(s)]$ , since the  $\rho$  resonance is the only one relevant in discussing  $\pi\pi$  threshold effects. The overall  $\text{Im}[\tilde{F}_{P\gamma^*\gamma}(M_V^2)]$  contribution would resemble thereby that in Fig. 3, but with additional sharp peaks at  $M_V = m_\omega, m_\phi$  locations, each of them weighted by the corresponding  $c_{\eta(\eta')V}$  coefficient. For  $M_V > m_P$ , the  $\gamma\gamma$  contribution dominates (the bluish region below  $m_\eta$  in Fig. 3 is almost negligible) and will be slightly modified when  $m_P > 2m_\pi$  due to the tail of the resonance contribution, whereas it will be less important when  $M_V < m_P$ .

<sup>11</sup>We explored further refined models with an improved threshold behavior for the  $\omega$  and  $\phi$  resonances. Given their narrow width they led to very similar results and we decided to take the ones in Eq. (4.3) for not obscuring our study and deviating the attention from our main concern, an estimation of a systematic error.



**Figure 3.** The imaginary part for integrand (4.5) expressed in terms of  $M_V^2 \text{Im}[K(M_V^2)]$  (black and dashed-purple lines for the  $\eta$  and  $\eta'$ , respectively) which has then to be convoluted with  $\text{Im}[\tilde{F}_{P\gamma^*\gamma}(M_V^2)]$ . As an illustration, we show  $\text{Im}[\tilde{F}_{P\gamma^*\gamma}(M_V^2)] = \text{Im}[G_\rho(M_V^2)]$  alone as a bluish area—the  $\text{Im}[G_{\omega,\phi}(M_V^2)]$  contributions would appear as sharp peaks at  $M_V = m_{\omega,\phi}$ .

Then, whenever  $\text{Im}[F_{P\gamma^*\gamma}(M_V^2 < m_P^2)] \neq 0$ ,  $\text{Im}[\mathcal{A}(m_P^2)]$  will be shifted with respect to the  $\gamma\gamma$  contribution. This is, as unitarity implies, whenever an intermediate hadronic channel appears below  $m_P$ .

For completeness, we illustrate in Table 2 the numerical shift in the imaginary part induced by the vector contributions with respect to the  $\gamma\gamma$  one using our toy model (4.1)—we stress that the three ( $V = \rho, \omega, \phi$ ) channels have been included in Table 2 calculations. The resulting shift illustrates the break of the unitary bound.

		$\gamma\gamma$	Total
$\text{Im}[\mathcal{A}_{\eta \rightarrow \ell\ell}(m_\eta^2)]$	$ee$	-21.920	-21.805
	$\mu\mu$	-5.468	-5.441
$\text{Im}[\mathcal{A}_{\eta' \rightarrow \ell\ell}(m_{\eta'}^2)]$	$ee$	-23.675	-19.251
	$\mu\mu$	-7.060	-5.733

**Table 2.** Imaginary part of  $\mathcal{A}(q^2)$  (total) compared to the imaginary part calculated from the  $\gamma\gamma$  channel alone. The hadronic contributions lower the total value of the imaginary part with respect to the  $\gamma\gamma$  contribution, invalidating the unitary bound.

Given that our model is a Stieltjes function, it is well known that the  $C_{N+1}^N(Q_1^2, Q_2^2)$  sequence is guaranteed to converge in the whole complex plane, except along the cut [40, 49], where zeros and poles of our CA will clutter to reproduce the discontinuity [40, 49, 54]<sup>12</sup>. Remarkably, even if the sequence does not converge along the cut, the integral along its imaginary part converges globally to that of the real function, which follows from Cauchy’s

<sup>12</sup>The reader may notice that, for the  $\eta$  case, the narrow  $\phi$  contribution produces a shift in sign near  $m_\phi$  for the imaginary part. We remark that in the limit of an infinitely narrow resonance, the resulting function is a combination of a Stieltjes and a meromorphic function, for which convergence applies as well [40]. Even if the particle would have a finite width, its effect would not affect convergence before the sign actually changes, which poses no problem for the  $\eta$  case.

$BR(P \rightarrow \ell\ell)$	toy model	$C_1^0$	Error (%)	$\mathcal{A}(m_P^2)$	toy model	$C_1^0$
$(\eta \rightarrow ee) \times 10^{-9}$	5.4095	5.4179	0.16	$(\eta \rightarrow ee)$	$31.4 - 21.8i$	$31.4 - 21.9i$
$(\eta \rightarrow \mu\mu) \times 10^{-6}$	4.49361	4.52701	0.74	$(\eta \rightarrow \mu\mu)$	$-1.09 - 5.44i$	$-1.05 - 5.47i$
$(\eta' \rightarrow ee) \times 10^{-10}$	1.70507	1.88331	9	$(\eta' \rightarrow ee)$	$46.4 - 19.2i$	$48.7 - 20.5i$
$(\eta' \rightarrow \mu\mu) \times 10^{-7}$	1.1953	1.46089	18	$(\eta' \rightarrow \mu\mu)$	$3.09 - 5.73i$	$3.82 - 6.10i$

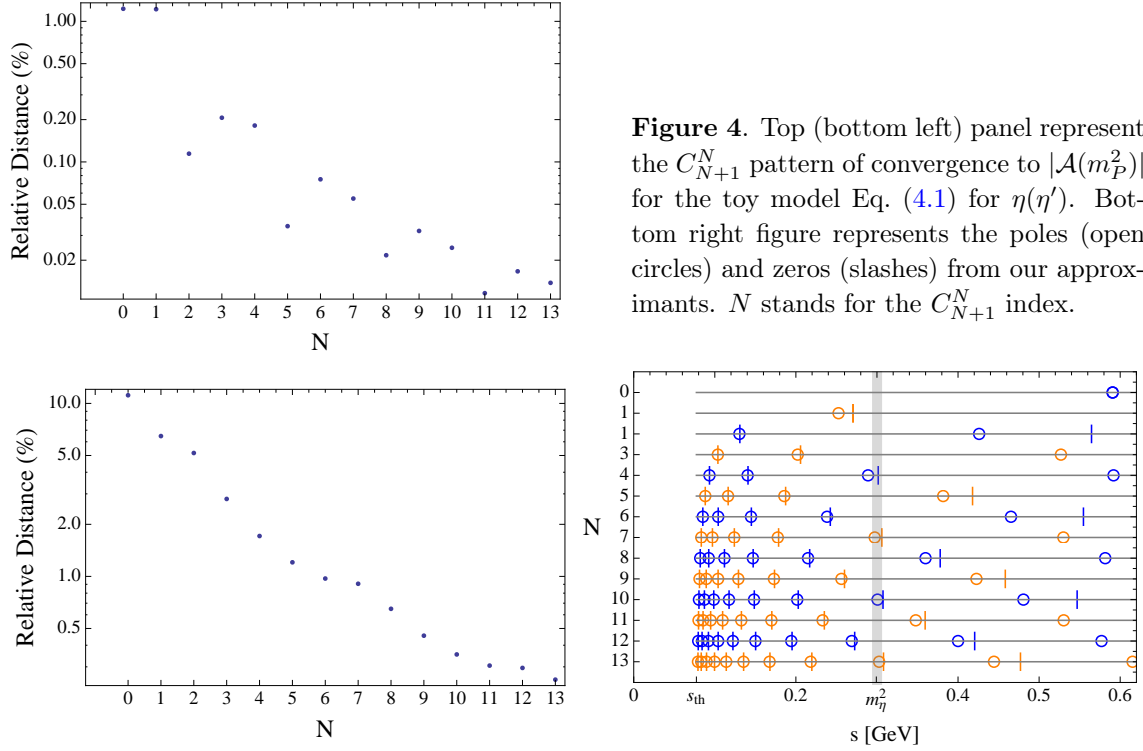
**Table 3.** Comparison between our toy model result and the simplest  $C_1^0(Q_1^2, Q_2^2)$  approximation for each channel. The Error represents the relative deviation between the model and the approximation. Left table collects the BR, whereas the right table contains the loop amplitude  $\mathcal{A}(m_P^2)$ . See details in the main text.

integral theorem. This means in practice that the approximant's poles will be responsible for effectively generating an imaginary part in our integral via the  $i\epsilon$  prescription mimicking the cut contribution. As an illustration, we collect the results for both BR and  $\mathcal{A}(m_P^2)$  from our simplest approximant, the  $C_1^0(Q_1^2, Q_2^2)$ , in Table 3 and compare its results with the toy model. The comparison of the BRs reveals a systematic error induced by the fact that we have truncated the CA sequence. For the  $\eta$ , such error is almost negligible (the role of the vector resonances is very mild there), whereas for the  $\eta'$  it goes almost up to 20%. These percentages will be used as an estimate of our systematic error in our final results for the  $C_1^0(Q_1^2, Q_2^2)$  element.

We would like to remark at this point that using a VMD model with the  $\rho$  mass — which was standard in the past for performing this calculation —, we would have found  $BR(\eta \rightarrow ee) = 5.30 \times 10^{-9}$ , which implies a larger systematic uncertainty compared to our results with the  $C_1^0(Q_1^2, Q_2^2)$  collected in Table 3, i.e.,  $BR(\eta \rightarrow ee) = 5.42 \times 10^{-9}$ . Using experimental data from the space-like region to fit the VMD model does not improve on the result. In this case, we would have obtained  $BR(\eta \rightarrow ee) = 5.26 \times 10^{-9}$ . These results illustrate the potential large systematic error coming from the usage of VMD data-fitting procedures from high energies for processes which are low-energy dominated, even if the quality of the fit is good enough and the errors tiny.

As we have said, the convergence of the CA sequence to our toy model is guaranteed in advance [40, 49, 54]. To fully show this feature beyond the simplest  $C_1^0(Q_1^2, Q_2^2)$  discussed in Table 3, we now discuss the the  $C_{N+1}^N(Q_1^2, Q_2^2)$  sequence up to  $N = 13$  and calculate the corresponding amplitude in Eq. (2.2). The relative distance in the complex plane of the energy squared, defined as  $|1 - \mathcal{A}(m_P^2)^{CA}/\mathcal{A}(m_P^2)|$ , is shown in Fig. 4, where, for simplicity, we employ only the  $G_\rho$  contribution in (4.1) (without  $\omega, \phi$  contributions).

The results in Fig. 4 show the ability of our approximants to systematically account for the TFF to arbitrary precision since the relative distance decreases when the order of the CA increases, even in the presence of the non-trivial behavior of the branch cut from the intermediate hadronic states. This exercise is a proof of concept that even though the CAs cannot reproduce the imaginary part locally, they are able to approximate it globally with great precision. Note the *a priori* irregular convergence for the  $\eta$  case in Fig. 4 (top panel). This is just an accident due to the combination of the particular TFF we are using together with the particular value of the  $\eta$  mass. Whenever some pole is located close to the  $\eta$  mass, it leads to a bad determination. This is compensated in higher approximants



**Figure 4.** Top (bottom left) panel represent the  $C_{N+1}^N$  pattern of convergence to  $|\mathcal{A}(m_P^2)|$  for the toy model Eq. (4.1) for  $\eta(\eta')$ . Bottom right figure represents the poles (open circles) and zeros (slashes) from our approximants.  $N$  stands for the  $C_{N+1}^N$  index.

with a nearby zero to this pole, alleviating this effect and making it negligible as  $N \rightarrow \infty$  as shown in Fig. 4 (bottom right panel), where the poles and zeros for different approximants are plotted.

We find then that, for the  $C_{N+1}^N(Q_1^2, Q_2^2)$  element, the systematic error can be accounted for, essentially, by the difference in the BR with respect to the  $C_N^{N-1}(Q_1^2, Q_2^2)$  result. As in our case of study we only reach the  $C_1^0(Q_1^2, Q_2^2)$  approximant, this procedure does not apply and we take as the systematic error for the BR the one which is displayed in the fourth column in Table 3 (based in our toy model). This possibly overestimates the systematic error, see comments in Sec. 5, but we opt for this to remain on the conservative side. As soon as experimental data on the double virtual TFF will be available, we will be able to extend our CA sequence and reduce the systematic error.

In Ref. [11] we considered a similar method for establishing the systematic error in the case of the  $\pi^0 \rightarrow e^+e^-$  case. There, we used the large- $N_c$  Regge model from Ref. [68] — which roughly factorizes at low energies, obeys the OPE, and for which  $a_{P;1,1} \gtrsim b_P^2$  — and obtained rather small ( $\mathcal{O}(1\%)$ ) systematic errors for the  $C_1^0(Q_1^2, Q_2^2)$  element, which was accounted for by the chosen  $a_{P;1,1}$  range. As remarked before, the large systematic error in Table 3 comes from the presence of the resonance region and the difficulty of approximating it with a single pole.

## 5 Final results

Our final results for the pseudoscalar decays into lepton pairs using the method here described have been calculated using Eq. (2.2) filled with the form factor data-driven param-

eterization described by Eq. (3.3).

The loop integral evaluation has been performed using `FeynCalc` [69] to obtain the decomposition into the Passarino-Veltman functions; the latter have been numerically evaluated using `LoopTools` [67]. In the factorization case ( $a_{P;1,1} = b_P^2$ ) this is possible using partial fraction decomposition, allowing to express the TFF and photon propagators in Eq. (2.2) as

$$\frac{\tilde{F}_{P\gamma^*\gamma^*}(k^2, (q-k)^2)}{k^2(q-k)^2} = \left( \frac{1}{k^2 - \frac{m_P^2}{b_P}} - \frac{1}{k^2} \right) \left( \frac{1}{(q-k)^2 - \frac{m_P^2}{b_P}} - \frac{1}{(q-k)^2} \right), \quad (5.1)$$

and the corresponding result involves one-, two- and three-point functions. For the OPE choice ( $a_{P;1,1} = 2b_P^2$ ) the loop integration is possible rewriting the TFF as

$$\tilde{F}_{P\gamma^*\gamma^*}(k^2, (q-k)^2) = \frac{-m_P^2/2b_P}{(k - \frac{q}{2})^2 - (\frac{m_P^2}{2b_P} - \frac{q^2}{4})}, \quad (5.2)$$

which allows to express the loop integral in terms of one-, two-, three- and four-point functions.

Later on, the branching ratios have been calculated with the formula given in Eq. (2.1). The results for the amplitude  $\mathcal{A}(m_P^2)$  and for the BRs are collected separately for illustrative purposes in Table 4, second column, and in Table 5, second column, respectively, and represent the main result of this work. In addition, we have included for completeness the  $Z$ -boson contribution (third column in Table 4)—we refer the interested reader to Section 7 or to Ref. [70]. Finally, these tables also include results when the kernel in Eq. (2.2) is expanded in terms of  $m_l/m_P$  as well as  $m_l/\Lambda$  and  $m_P/\Lambda$ , with  $\Lambda$  a cut-off of the loop integral (fourth column in Tables 4 and 5).

To report a final result, several sources of errors must be taken into account:

- Our ignorance on the value of the slope of double virtuality  $a_{P;11}$ , Eq. (3.3), is captured by the range  $a_{P;11} \in (b_P^2 \div 2b_P^2)$ . With this range, we obtain the results for the loop amplitude  $\mathcal{A}(m_P^2)$ , Eq. (2.2), collected in Table 4, and its associated BR displayed in Table 5. The values, which are reported in a range form corresponding to imposing the OPE or the factorization, include statistical errors for the amplitudes and statistical and systematic errors for the branching ratios<sup>13</sup>.
- The single error given for the amplitude results in the second column of Table 4 corresponds to the error of the parameter  $b_P$  in Eq. (2.1) used for the TFF reconstruction. This parameter already includes a systematic error on top of the statistical one, which was estimated in Refs. [55, 56] for  $\eta$  and  $\eta'$ , respectively.
- The results collected in Table 5 contain three different sources of error. The first corresponds to the often neglected uncertainty from  $\text{BR}(P \rightarrow \gamma\gamma)$  [6] in Eq. (2.1). This error is relevant for the  $\eta$  and very important for the  $\eta'$ . The second source is

---

<sup>13</sup>The systematic error shown in Table 3 (column called Error) and discussed in Sec. 4 is only intended for the BR rather than for the amplitude. This is, however, enough for our purposes.



Process	$\mathcal{A}(m_P^2)$	$\mathcal{A}^Z(m_P^2)$	$\mathcal{A}^{\text{app}}(m_P^2)$
$\eta \rightarrow e^+e^-$	$(30.95 \div 31.51)(11) - 21.92(0)i$	$-0.03$	$(27.53 \div 28.00) - 21.92i$
$\eta \rightarrow \mu^+\mu^-$	$-(1.52 \div 0.99)(5) - 5.47(0)i$	$-0.03$	$-(2.33 \div 1.87) - 5.47i$
$\eta' \rightarrow e^+e^-$	$(47.4 \div 48.2)(5) - (21.0)(5)i$	$0.03$	$(35.20 \div 35.66) - 23.68i$
$\eta' \rightarrow \mu^+\mu^-$	$(2.95 \div 3.65)(19) - (6.27)(17)i$	$0.03$	$-(0.66 \div 0.20) - 7.06i$
$\pi^0 \rightarrow e^+e^-$	$(10.00 \div 10.46)(12) - 17.52i$	$-0.05$	$(9.84 \div 10.30) - 17.52i$

**Table 4.** Our results for the range  $a_{P;11} \in (2b_P^2 \div b_P^2)$ , corresponding to (OPE÷Factorization). The error refers to the statistical error alone. We quote the  $Z$ -boson contribution  $\mathcal{A}^Z(m_P^2)$  separately and quote the approximated  $\mathcal{A}^{\text{app}}(m_P^2)$  calculation after expanding Eq. (2.2) in terms of  $m_l/m_P$  as well as  $m_l/\Lambda$  and  $m_P/\Lambda$ . See details in the main text.

Process	BR	BR w/Z	BR app
$\eta \rightarrow e^+e^-$	$(5.31 \div 5.44)(3)(2)(1) \times 10^{-9}$	$(5.32 \div 5.45) \times 10^{-9}$	$(4.58 \div 4.68) \times 10^{-9}$
$\eta \rightarrow \mu^+\mu^-$	$(4.72 \div 4.52)(2)(3)(4) \times 10^{-6}$	$(4.70 \div 4.51) \times 10^{-6}$	$(5.16 \div 4.88) \times 10^{-6}$
$\eta' \rightarrow e^+e^-$	$(1.82 \div 1.87)(7)(2)(16) \times 10^{-10}$	$(1.82 \div 1.87) \times 10^{-10}$	$(1.22 \div 1.24) \times 10^{-10}$
$\eta' \rightarrow \mu^+\mu^-$	$(1.36 \div 1.49)(5)(3)(25) \times 10^{-7}$	$(1.35 \div 1.48) \times 10^{-7}$	$(1.42 \div 1.41) \times 10^{-7}$
$\pi^0 \rightarrow e^+e^-$	$(6.20 \div 6.35)(0)(4)(1) \times 10^{-8}$	$(6.22 \div 6.36) \times 10^{-8}$	$(6.17 \div 6.31) \times 10^{-8}$

**Table 5.** Our results for the Branching Ratios for the range  $a_{P;11} \in (2b_P^2 \div b_P^2)$ , corresponding to (OPE÷Factorization). The errors refers to the statistical error for  $\text{BR}(P \rightarrow \gamma\gamma)$ , the error from  $b_P$  and the systematic, respectively. We compare to the results either neglecting the  $Z$ -boson contribution (BR w/Z) or after expanding Eq. (2.2) in terms of  $m_l/m_P$  as well as  $m_l/\Lambda$  and  $m_P/\Lambda$  (BR app) discussed in the main text.

the one coming from  $b_P$  as discussed in the previous item. The third is the systematic error from our method discussed in the previous section and given in Table 3, fourth column.

For completeness, Tables 4 and 5 also include the results for the  $\pi^0 \rightarrow e^+e^-$  process [11].

From those values it is clear that, for the  $\eta$  decays, the most pressing task is the improvement of the double-virtual description which currently limits the theoretical uncertainty (the difference in the BRs between OPE and factorization is of about  $0.1 \times 10^{-9}$  for the  $ee$  channel, and about  $0.2 \times 10^{-6}$  for the  $\mu\mu$  channel, far beyond the impact of the errors quoted). Notice that the systematic error of our method is of the order of the other error sources. To reduce our ignorance on the double virtual TFF, experimental data (or, eventually, lattice simulations) on the double-virtual TFF would be required.

Concerning the  $\eta'$ , the situation is more complicated. In Table 5, third row, the largest errors arise from our systematic-error estimation in Table 3, while the errors coming from the normalization or the slope of the TFF are milder. The systematic uncertainty could be dramatically reduced if instead of estimating it using the toy model in Sec. 4, we would look at the difference between the  $C_1^0(Q_1^2, Q_2^2)$  and the  $C_2^1(Q_1^2, Q_2^2)$  elements in the factorization approach (where no knowledge of the double virtuality is required). This is not surprising since, as emphasized previously, our toy model is not realistic enough and fails to describe the space-like region, which may hint an overestimation of the systematic uncertainty.

In this respect, a more elaborated investigation including not only the low- and high-energy behaviors, but the information about the time-like region, such as physical res-



onances and threshold discontinuities, would be of interest in order to reach a similar precision to what is achieved in the  $\eta$  case. Investigations in this respect are undergoing.

Our results may be compared to the experimental values given in Table 1. We find an interesting deviation in the  $\eta \rightarrow \mu^+\mu^-$  channel. Still, the experimental accuracy prevents us from drawing any conclusion and a new experiment would be very welcomed. It has been recently suggested that this could be possible at the LHCb Collaboration [23]. The result becomes even more interesting when comparing to the analogous  $\pi^0 \rightarrow e^+e^-$  anomaly (cf. our results in Table 5 and the experimental value in Table 1), as we find that, whereas for the  $\pi^0$  case a very damped TFF at large energies was required to reproduce the experimental value [11], the  $\eta$  case demands a smoothly falling TFF instead, which points to a puzzling situation. Very interesting as well is the current bound on  $\eta' \rightarrow e^+e^-$ , which is getting closer to the theoretical expectations. In this respect, it would be very stimulating to push for a new measurement. Such an effort is currently ongoing at VEPP-2000  $e^+e^-$  collider at Novosibirsk, where they plan to increase their statistics by a factor of ten. Additionally, it has been suggested the possibility to measure  $\eta' \rightarrow \mu^+\mu^-$  at LHCb [23].

The results above represent an important improvement with respect to previous studies. See for instance the comparison in Tables 4 and 5 between our exact results (second column) and the results obtained (last column) if we would have used the standard approximated calculation, widely used in the literature [26, 35, 42], which amounts to expand the kernel in Eq. (2.2) in terms of  $m_l/m_P$  as well as  $m_l/\Lambda$  and  $m_P/\Lambda$ , with  $\Lambda$  a cut-off of the loop integral (the hadronic scale driving the TFF), see Eq. (3) in Ref. [11]. Indeed, the missing corrections concerning the lepton mass along the lines of Ref. [36] are relevant for the  $\mu^+\mu^-$  channel, which regrettably were not included in the recent Ref. [10].

Equally important is the correct implementation of both low- and high-energy TFF's behavior (i.e. using  $b_P$  instead of the commonly employed resonance or VMD fit parameters as well as accounting for the OPE behavior), which for the  $\eta$  case induces the largest effect. Beyond that, the systematic error estimation is for the first time discussed and is by no means negligible.

On the contrary, the role of the  $Z$  boson is almost negligible, though similar to some hadronic uncertainties. This is already a sign that new physics at the electroweak scale will not alter the results here presented and any eventual deviation from the Standard Model in these decays will demand a compelling New Physics benchmark.

## 6 The low-energy description: the $\chi$ PT approach

After reporting on our final results, we would like to discuss the comments raised in the Introduction about the calculation of the pseudoscalar decays into lepton pairs within the framework of  $\chi$ PT. In this framework, the value for  $\mathcal{A}(m_P^2)$  at leading order is given, in the  $\overline{MS}$  renormalization scheme, as [30, 34]

$$\mathcal{A}^{\text{LO}}(m_P^2) = \frac{i\pi}{2\beta_\ell} L + \frac{1}{\beta_\ell} \left[ \frac{1}{4} L^2 + \frac{\pi^2}{12} + Li_2 \left( \frac{\beta_\ell - 1}{1 + \beta_\ell} \right) \right] - \frac{5}{2} + \frac{3}{2} \ln \left( \frac{m_\ell^2}{\mu^2} \right) + \chi(\mu), \quad (6.1)$$

	$\pi^0 \rightarrow e^+e^-$	$\eta \rightarrow e^+e^-$	$\eta \rightarrow \mu^+\mu^-$	$\eta' \rightarrow e^+e^-$	$\eta' \rightarrow \mu^+\mu^-$
$\chi(\mu)$	$(2.53 \div 2.99)$	$(5.90 \div 6.46)$	$(3.29 \div 3.82)$	$(14.2 \div 14.9) + 2.52i$	$(5.61 \div 6.31) + 0.75i$
$\chi(\mu)_{m_\pi}$	$(2.53 \div 2.99)$	$(2.66 \div 3.12)$	—	$(2.16 \div 2.62)$	—
$\chi(\mu)_{UV}$	$(2.53 \div 2.99)$	$(5.50 \div 6.05)$	$(3.11 \div 3.64)$	$(16.8 \div 17.7) + 7.09i$	$(6.56 \div 7.35) + 2.12i$

**Table 6.** Comparison between  $\chi$ PT counter-term  $\chi(\mu)$  defined in Eq. (6.1) with its equal-mass version,  $\chi(\mu)_{m_\pi}$ , and the  $U(3)$ -symmetric version  $\chi(\mu)_{UV}$ . Results are always reported for the range (OPE $\div$ Fact). See description in the main text.

where  $L = \ln\left(\frac{1-\beta_\ell}{1+\beta_\ell}\right)$ ,  $\beta_\ell$  is defined below Eq. (2.1),  $\mu$  stands for the renormalization scale and  $\chi(\mu) \equiv -(\chi(\mu)_1 + \chi(\mu)_2)/4$  is the (scale-dependent) counter-term. Given that, at leading order in  $\chi$ PT,  $\chi(\mu)$  is common to all the  $P \rightarrow \bar{\ell}\ell$  decays<sup>14</sup>, it has been argued [31] that any of the experimental values for these decays may be used to predict the rest.

To illustrate the accuracy of  $\chi$ PT at leading order, we want to quantify the role of the different approximations within  $\chi$ PT when calculating the parameter  $\chi(\mu)$  from Eq. (6.1). For that purpose, we equate our final results for  $\mathcal{A}(m_P^2)$  (see Table 4) to the  $\chi$ PT result, Eq. (6.1), from which  $\chi(\mu)$  can be determined. The values obtained are displayed as  $\chi(\mu)$  in Table 6 for  $\mu = 0.77$  GeV and represent the second main result in this work. In each entry on this table, we quote the extreme solutions given by the boundaries of the (OPE $\div$ Fact) choices for  $a_{P;1,1}$ . The spread of values obtained for  $\chi(\mu)$  in each channel manifest that the  $\chi$ PT predictions at this order are subject to significant corrections, which origin is detailed below.

First, there is a large splitting due to  $U(3)$ -breaking effects coming from the fact that  $m_\pi < m_\eta < m_{\eta'}$ . To see the magnitude of them, we illustrate in Table 6, row  $\chi(\mu)_{m_\pi}$ , what would have been obtained for  $\chi(\mu)$  if we would have imposed  $m_{\eta'} = m_\eta = m_\pi \equiv m_\pi$  in our  $\mathcal{A}(m_P^2)$  calculations as well as in Eq. (6.1). In this scenario, since  $m_\pi < 2m_\mu$ , the decays into muon pairs cannot be calculated and are accounted for by a hyphen.

An additional difference arises from the different pseudoscalar meson transition form factors. Imposing that all of them are equal to the  $\pi^0$  TFF, but the masses are restored to their real values, the results are collected under the row called  $\chi(\mu)_{UV}$  in Table 6.

Finally,  $m_\ell$  plays as well a relevant role in our decays. Corrections of the order  $\mathcal{O}(m_\ell/m_P, \Lambda)$  are expected to arise dynamically, for instance, through pion loop effects as in Fig. 2 and implies that calculations at the next-to-leading order are relevant in  $\chi$ PT. This is specially important for the  $\eta'$ , as this is the only mechanism able to generate an additional imaginary contribution within the  $\chi$ PT framework. Moreover, this puts a word of caution on naive  $\chi$ PT analysis of these processes when looking for New Physics, as a different experimental extraction for  $\chi(\mu)$  in either the electronic or muonic channel (for the same pseudoscalar) is expected within the Standard Model and does not necessarily imply lepton flavor violation.

<sup>14</sup>This actually requires some assumptions when including the  $\eta'$ , or assumptions to avoid the  $\eta - \eta'$  mixing if this is not included—which is known to be extremely important to get the  $P \rightarrow \gamma\gamma$  decays right. Invoking further assumptions, it has been shown that  $K_L^0 \rightarrow \bar{\ell}\ell$  decays could be used as well [31, 34].

## 6.1 The $\pi^0$ -exchange contribution to the $2S$ hyperfine-splitting in the muonic hydrogen

The results collected in Table 6, first row, are also relevant for calculating the  $\pi^0$ -exchange contribution to the  $2S$  hyperfine-splitting in the muonic hydrogen [71]. Such calculation can be performed within  $\chi$ PT [72], which involves again Eq. (6.1). However, the kinematics of the process involve a vanishingly small  $Q^2$  space-like momentum for the  $\pi^0$  since the  $\pi^0$ -exchange contribution to the  $2S$  hyperfine-splitting is a  $t$ -channel exchange. As such, it is  $\mathcal{A}(Q^2 \simeq 0)$  instead of  $\mathcal{A}(m_\pi^2)$  which is relevant now [71], shifting the values obtained in Table 6. To illustrate this, we recalculate  $\mathcal{A}(Q^2)$  from Eq. (2.2) taking the limit  $Q^2 \rightarrow 0$ , and obtain the new subtraction constant which should be used in Eq. (6.1) to reproduce our results. Using the  $\pi^0$  TFF [51] we obtain

$$\chi_{\pi^0}^{ee}(\mu) = (2.37 \div 2.83) \quad (6.2)$$

for  $\ell = e$ , which is smaller than its counterpart collected in Table 6. However, for the  $2S$  hyperfine-splitting in muonic hydrogen what is needed is the coupling to muons ( $\ell = \mu$ ). In that case, we obtain<sup>15</sup>

$$\chi_{\pi^0}^{\mu\mu}(\mu) = (2.18 \div 2.63), \quad (6.3)$$

which is even lower than Eq. (6.2). Note that the shift is of the order of the uncertainties quoted in Table 4 and arises again from the full  $q^2$  and  $m_\ell^2$  dependence in Eq. (2.2), which is not accounted for at LO in  $\chi$ PT. We note that the relation between the  $\chi_{\pi^0}^{ee}(\mu)$  in Eq. (6.2) and  $\chi_{\pi^0}^{\mu\mu}(\mu)$  in (6.3) and the one extracted from the experimental results is non-trivial as it is TFF dependent. In quoting our results, we implicitly assume that there is no New-Physics contribution. However, if the current discrepancies among theory and experiment persists, indicating New Physics contribution—which we will discuss in Sec. 7—the connection between the experimental  $\chi(\mu)$  and that in Eqs. (6.2) and (6.3) will depend on the particular scenario and will have to be reanalyzed.

The results above are illustrative as well regarding  $(g - 2)_\mu$  hadronic contributions, which in  $\chi$ PT involve  $\chi(\mu)$  together with an additional counterterm,  $C(\mu)$ , as an input [73]. If we were able to determine  $C(\mu)$  somehow, from  $(g - 2)_e$  for example, and  $\chi(\mu)$  would be taken from the experimental  $\pi^0 \rightarrow e^+e^-$  result, extrapolating up to the  $\mu$  case may imply a non-negligible error as illustrated above; similar effects may arise for  $C(\mu)$  as well [11, 74].

## 6.2 Higher order corrections

As discussed above, the precision which is reached at the LO in  $\chi$ PT for processes involving a  $P\bar{\ell}\ell$  vertex may not be enough—a feature which manifests when comparing the same process for a different  $\ell = e, \mu$  channel. This suggests to look at the next-to-leading order.

---

<sup>15</sup>Our results Eqs. (6.2) and (6.3) do not change to the quoted digits when extrapolating to  $\mathcal{A}(Q^2)$  up to energies  $Q^2 \sim (50)^2 \text{ MeV}^2$ , roughly at distances below 1.5% the muonic Bohr radius.

In this respect,  $\chi$ PT would yield a power series expansion for the TFF<sup>16</sup>

$$\tilde{F}_{P\gamma^*\gamma^*}(q_1^2, q_2^2) = \underbrace{1}_{\text{LO}} + \underbrace{\frac{1}{\Lambda^2}(q_1^2 + q_2^2)}_{\text{NLO}} + \underbrace{\frac{1}{\Lambda^4}(q_1^4 + q_2^4) + \frac{1}{\Lambda^4}(q_1^2 q_2^2)}_{\text{NNLO}} + \mathcal{O}\left(\frac{q^6}{\Lambda^6}\right). \quad (6.4)$$

Then, we could calculate the result of (2.2) for the TFF in (6.4),

$$\begin{aligned} \mathcal{A}(q^2, m_\ell^2) &= \frac{2i}{\pi^2 q^2} \int d^4 k \frac{(k^2 q^2 - (k \cdot q)^2)}{k^2 (q - k)^2 ((p - k)^2 - m_\ell^2)} \left[ 1 + \frac{(\dots)}{\Lambda^2} + \frac{(\dots)}{\Lambda^4} + \dots \right] \\ &\equiv \mathcal{A}^{\text{LO}}(q^2, m_\ell^2) + \mathcal{A}^{\text{NLO}}(q^2, m_\ell^2) + \mathcal{A}^{\text{NNLO}}(q^2, m_\ell^2) + \dots, \end{aligned} \quad (6.5)$$

where  $\mathcal{A}^{\text{LO}}(q^2)$  has been given in (6.1) and

$$\mathcal{A}^{\text{NLO}}(q^2, m_\ell^2) = \frac{1}{3\Lambda^2}(q^2 - 10m_\ell^2)(1 - L_\ell) + \frac{1}{9\Lambda^2}(4m_\ell^2 - q^2), \quad (6.6)$$

$$\mathcal{A}^{\text{NNLO}}(q^2, m_\ell^2) = \left[ \frac{126m_\ell^4 - q^4 - 8m_\ell^2 q^2}{12\Lambda^4} L_\ell + \frac{26m_\ell^2 q^2 + 7q^4 - 702m_\ell^4}{72\Lambda^4} \right], \quad (6.7)$$

where  $L_\ell = \ln(m_\ell^2/\Lambda^2)$ <sup>17</sup>.

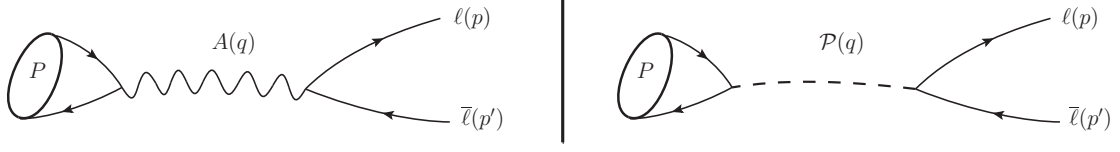
We notice that the LO leading logs  $L_\ell$  correspond —not surprisingly as they arise from a power-like expansion as well— to the corrections found in [9, 10] if  $\Lambda$  is taken as the VMD scale. We adopt then a much more modest approach and retain the leading logs alone, which represents a good approximation. This would produce a straightforward generalization to higher orders as well as a tool to estimate the convergence of the chiral expansion. Of particular relevance is the difference  $\mathcal{A}(q^2, m_e^2) - \mathcal{A}(q^2, m_\mu^2)$ , where one expects a better convergence of (6.5) due to partial cancellations. Taking into account the smallness of the electron mass, we find that such a shift is given as

$$\begin{aligned} \mathcal{A}(q^2, m_e^2) - \mathcal{A}(q^2, m_\mu^2) &= \mathcal{A}^{\text{LO}}(q^2, m_e^2) - \mathcal{A}^{\text{LO}}(q^2, m_\mu^2) \\ &\quad + \frac{q^2}{3\Lambda^2} \left( 1 + \frac{q^2}{4\Lambda^2} \right) \ln \left( \frac{m_\mu^2}{m_e^2} \right) + \frac{10m_\mu^2}{3\Lambda^2} \ln \left( \frac{\Lambda^2}{m_\mu^2} \right). \end{aligned} \quad (6.8)$$

Whereas our theoretical results for the leptonic and muonic channels in Tab. 4 could not be reproduced at LO with an unique counterterm, the observed differences in Tab. 6 and Section 6.1 can be easily accounted for to a good approximation taking into account the additional terms in (6.8) —an exception is the  $\eta'$  case, for which the pion loops cannot be neglected in order to extract an imaginary part. The expansion above, Eq. (6.8), proves extremely useful to relate different leptonic channels, which is not only relevant in the cases discussed above, but for  $\chi$ PT studies on lepton flavor violation in  $K_L \rightarrow \bar{\ell}\ell$  decays [75].

<sup>16</sup>For simplicity, we have assumed a single scale for the TFF inspired in typical VMD models. Note that logarithmic terms coming from loops are of course present too. However, they are subleading as compared to the power expansion and may be Taylor expanded for the  $\pi^0$  and  $\eta$  cases.

<sup>17</sup>We note that, previous to the renormalization procedure, the loop integral produces  $L_\ell = \ln(m_\ell^2/\mu^2)$  terms. It is after including the (non-explicitly shown) counterterms of the theory that the (scale independent)  $L_\ell = \ln(m_\ell^2/\Lambda^2)$  result would appear. We emphasize that it is not our aim to perform a detailed higher-order  $\chi$ PT evaluation, rather than to illustrate the effects which are expected to appear.



**Figure 5.** Left(right): additional tree level contributions from an axial(pseudoscalar) field. The  $P$  within the blob stand for the pseudoscalar meson;  $A(\mathcal{P})$  stands for the axial(pseudoscalar) field with momentum  $q$ ;  $\ell(\bar{\ell})$  for the (anti)lepton with momentum  $p(p')$ .

## 7 New physics contributions

At this point, we are finally on a firm foot to discuss about possible New Physics (NP) contributions given the current discrepancies in the two existing measured decays. As discussed in [11], any additional contribution will always manifest, after Fierz-rearrangement, only through effective pseudoscalar ( $\mathcal{P}$ ) and axial ( $A$ ) contributions, which given the existing well-motivated models, are conveniently expressed in the effective Lagrangian

$$\mathcal{L} = \frac{g}{4m_W} \sum_f m_A c_f^A (\bar{f} \not{A} \gamma_5 f) + 2m_f c_f^{\mathcal{P}} (\bar{f} i \gamma_5 f) \mathcal{P},$$

where  $g, m_W$  are the standard electroweak parameters, and  $c_f^{A, \mathcal{P}}$  are dimensionless couplings to the fermions  $f = \{u, d, s, e, \mu\}$ . These interactions yield additional tree-level contributions as shown in Fig. 5. Their corresponding amplitudes read

$$i\mathcal{M} = \frac{igc_\ell^A m_A}{4m_W} [\bar{u}_{p,s} \gamma_\mu \gamma_5 v_{p',s'}] \frac{-i \left( g_{\mu\nu} - \frac{q_\mu q_\nu}{m_A^2} \right)}{m_P^2 - m_A^2} \frac{igm_A}{4m_W} \overbrace{\sum_q \langle 0 | c_q^A \bar{q} \gamma^\mu \gamma_5 q | P(q) \rangle}^{\langle 0 | J_{\mu 5}^{\text{NP}} | P(q) \rangle}, \quad (7.1)$$

$$i\mathcal{M} = \frac{igc_\ell^{\mathcal{P}}}{2m_W} m_\ell [\bar{u}_{p,s} i \gamma_5 v_{p',s'}] \frac{i}{m_P^2 - M_{\mathcal{P}}^2} \frac{ig}{2m_W} \overbrace{\sum_q \langle 0 | c_q^{\mathcal{P}} m_q \bar{q} i \gamma_5 q | P(q) \rangle}^{\langle 0 | \mathcal{P}^{\text{NP}} | P(q) \rangle}, \quad (7.2)$$

for the axial and pseudoscalar contribution with masses  $m_A$  and  $M_{\mathcal{P}}$ , respectively.

The hadronic matrix element  $\langle 0 | J_{\mu 5}^{\text{NP}} | P(q) \rangle$  in Eq. (7.1) may be easily obtained using the octet-singlet basis or the flavor basis as an analogy to the SM axial current. From the pseudoscalar decay constants, defined in terms of the  $U(3)$  axial currents<sup>18</sup>,

$$\langle 0 | J_{\mu 5}^a | P(q) \rangle \equiv i q_\mu F_P^a, \quad J_{\mu 5}^a = \bar{q} \gamma_\mu \gamma_5 \frac{\lambda^a}{2} q, \quad q = (u, d, s)^T,$$

where  $a = 8, 1$  and  $P = \eta, \eta'$ , we find that

$$\begin{aligned} \langle 0 | J_{\mu 5}^{\text{NP}} | P(p) \rangle &= \sum_a \langle 0 | \text{Tr}(J_{\mu 5}^{\text{NP}} \lambda^a) J_{\mu 5}^a | P(p) \rangle \\ &= \sum_a \text{Tr}(\text{diag}(c_u^A, c_d^A, c_s^A) \lambda^a) \langle 0 | J_{\mu 5}^a | P(p) \rangle. \end{aligned}$$

<sup>18</sup>The singlet matrix is defined as  $\lambda^1 = \sqrt{2/3} \mathbb{1}$ .

Inserting this equation back into Eq. (7.1) and using the spinor projector for the singlet state from Ref. [76], we obtain

$$\begin{aligned} i\mathcal{M} &= -ic_\ell^A \frac{G_F}{\sqrt{2}} m_\ell (\bar{u}_{p,s} i\gamma_5 v_{p',s'}) \sum_a \text{Tr}(J_{\mu 5}^{\text{NP}} \lambda^a) F_P^a \\ &= c_\ell^A m_\ell m_P G_F \sum_a \text{Tr}(J_{\mu 5}^{\text{NP}} \lambda^a) F_P^a, \end{aligned}$$

where  $G_F = \frac{\sqrt{2}g^2}{8m_W^2}$ . This produces an additional contribution to Eq. (2.2) which reads

$$\mathcal{A}(q^2) \rightarrow \mathcal{A}(q^2) + \frac{\sqrt{2}G_F}{4\alpha^2 F_{P\gamma\gamma}} c_\ell^A \sum_a \text{Tr}(J_{\mu 5}^{\text{NP}} \lambda^a) F_P^a.$$

As an example, the  $Z$ -boson contribution is obtained after taking  $c_u^Z = -c_{d,s,e,\mu}^Z = 1$ , leading for  $P = \{\pi^0, \eta, \eta'\}$

$$\mathcal{A}(q^2) \rightarrow \mathcal{A}(q^2) - \frac{2\sqrt{2}G_F F_\pi}{4\alpha^2 F_{P\gamma\gamma}} \left\{ 1, \frac{F_\eta^8}{\sqrt{3}F_\pi} - \frac{F_\eta^1}{\sqrt{6}F_\pi}, \frac{F_{\eta'}^8}{\sqrt{3}F_\pi} - \frac{F_{\eta'}^1}{\sqrt{6}F_\pi} \right\}.$$

Alternatively, we could have used the flavor basis instead [56]<sup>19</sup>, then,

$$\mathcal{A}(q^2) \rightarrow \mathcal{A}(q^2) - \frac{2\sqrt{2}G_F F_\pi}{4\alpha^2 F_{P\gamma\gamma}} \left\{ 1, -\frac{F_\eta^s}{\sqrt{2}F_\pi}, -\frac{F_{\eta'}^s}{\sqrt{2}F_\pi} \right\}.$$

For the pseudoscalar contribution, the  $\langle 0 | \mathcal{P}^{\text{NP}} | P(q) \rangle$  hadronic matrix element determination in Eq. (7.2) is more complicated whenever the singlet component is involved. This is the case for both  $\eta$  and  $\eta'$  as they are an admixture of the octet and singlet states. To illustrate this, let us calculate such matrix elements at leading order (LO) in  $\chi\text{PT}$ , which amounts to retain the leading pseudoscalar  $\mathcal{P} \times P$  term arising from the interaction between the pseudoscalar field  $P$  and the pseudoscalar current  $\mathcal{P}$  defined in  $\chi\text{PT}$  from the building block  $\chi \equiv 2B_0 i\mathcal{P}$  ( $B_0$  is the low-energy constant related to the scalar singlet quark condensate  $\langle \bar{q}q \rangle_0$  in the chiral limit). Then, in the presence of NP of pseudoscalar type,  $\chi \rightarrow 2B_0 i\mathcal{P}^{\text{NP}}$ . For the  $\pi^0$  such term corresponds to

$$F_0 B_0 \hat{m} (c_u^{\mathcal{P}} - c_d^{\mathcal{P}}) \mathcal{P}^{\text{NP}} \pi^0,$$

with  $\hat{m} = (m_u + m_d)/2$ , from which the matrix element reads ( $2B_0 \hat{m} = m_\pi^2$ )

$$\langle 0 | \mathcal{P}^{\text{NP}} | \pi^0 \rangle = F_0 B_0 \hat{m} (c_u^{\mathcal{P}} - c_d^{\mathcal{P}}) = \frac{F_\pi}{2} m_\pi^2 (c_u^{\mathcal{P}} - c_d^{\mathcal{P}}).$$

For the  $\eta$  and  $\eta'$  such term is more involved and at LO reads

$$F_0 B_0 \mathcal{P}^{\text{NP}} \left( \frac{1}{\sqrt{3}} (\hat{m}(c_u^{\mathcal{P}} + c_d^{\mathcal{P}}) - 2c_s^{\mathcal{P}} m_s) \eta_8 + \sqrt{\frac{2}{3}} (\hat{m}(c_u^{\mathcal{P}} + c_d^{\mathcal{P}}) + c_s^{\mathcal{P}} m_s) \eta_1 \right).$$

---

<sup>19</sup>Which amounts to trade  $\lambda^8$  and  $\lambda^1$  for  $\lambda^q = \text{diag}(1, 1, 0)$  and  $\lambda^s = \text{diag}(0, 0, \sqrt{2})$ .

After relabeling, introducing  $g_8 \equiv (c_u^{\mathcal{P}} + c_d^{\mathcal{P}} - 2c_s^{\mathcal{P}})/\sqrt{3}$  and  $g_1 \equiv \sqrt{2}(c_u^{\mathcal{P}} + c_d^{\mathcal{P}} + c_s^{\mathcal{P}})/\sqrt{3}$ , it can be expressed as

$$\frac{F_0}{2} (\eta_8(g_8 M_8^2 + g_1 M_{81}^2) + \eta_1(g_8 M_{81}^2 + g_1 M_1^2)) \mathcal{P}^{\text{NP}},$$

where  $M_8^2$ ,  $M_1^2$  and  $M_{81}^2$  are defined in Refs. [77, 78]. Finally, using the  $\eta - \eta'$  masses, mixing and decay constants at LO<sup>20</sup>, we obtain for the matrix element

$$\langle 0 | \mathcal{P}^{\text{NP}} | \eta(\eta') \rangle = \sum_a \frac{1}{2} F_{\eta(\eta')}^a g_a m_{\eta(\eta')}^2 \left( 1 - \delta^{a1} \frac{M_0^2}{M_{\eta(\eta')}^2} \right), \quad (7.3)$$

where  $g_a$  has been defined above and  $M_0^2 = 6\tau/F_0^2$  is the topological mass term [78]. After some algebra, we have obtained a relation which is very similar to the  $\pi^0$  result —except for the singlet  $a = 1$  term— and resembling that of the axial current matrix element. The previous calculation applies so far up to LO. However, such precision is not enough to reproduce the observed  $\eta - \eta'$  mixing, which requires higher order calculations. In order to provide a general result, valid at any order and related to known parameters, it is convenient to recall the Ward identity [78]

$$\partial^\mu J_{\mu 5}^a = \{\mathcal{P}^a, \mathcal{M}\} + \delta^{a1} \sqrt{N_F/2} \omega; \quad \mathcal{P}^a = \bar{q} i \gamma_5 \frac{\lambda^a}{2} q, \quad q = (u, d, s)^T,$$

where  $\mathcal{M} = \text{diag}(\hat{m}, \hat{m}, m_s)$  is the quark mass matrix. In such a way, the pseudoscalar current may be expressed in terms of the axial current and the winding number density  $\omega$ . Then, using the same algebra as previously, the matrix element can be expressed as

$$\begin{aligned} \langle 0 | \mathcal{P}^{\text{NP}} | P(p) \rangle &= \frac{1}{2} \sum_a \text{Tr}(\mathcal{P}^{\text{NP}} \lambda^a) \langle 0 | \partial^\mu J_{\mu 5}^a - \delta^{a1} \sqrt{3/2} \omega | P(p) \rangle \\ &= \frac{m_P^2}{2} \sum_a \text{Tr}(\text{diag}(c_u^{\mathcal{P}}, c_d^{\mathcal{P}}, c_s^{\mathcal{P}}) \lambda^a) F_P^a (1 - \Delta \delta^{1a}), \end{aligned} \quad (7.4)$$

where  $\Delta = \langle 0 | \sqrt{3/2} \omega | P \rangle / m_P^2 F_P^1$ . Still,  $\Delta$  needs to be determined. Neglecting the  $u$  and  $d$  quark masses,  $\omega$  may be expressed as [79]:

$$\sqrt{3/2} \omega = \partial^\mu J_{\mu 5}^1 + \frac{1}{\sqrt{2}} \partial^\mu J_{\mu 5}^8.$$

Plugging this relation into (7.4), we obtain  $\Delta = 1 + F_P^8/(\sqrt{2} F_P^1)$ , so the pseudoscalar contribution to  $P \rightarrow \bar{\ell} \ell$  can be finally expressed as

$$\begin{aligned} i\mathcal{M} &= -i[\bar{u}_{p,s} i \gamma_5 v_{p',s'}] \frac{G_F}{\sqrt{2}} \frac{m_P^2 m_\ell c_\ell^{\mathcal{P}}}{m_P^2 - M_P^2} \sum_a \text{Tr}(\mathcal{P}^{\text{NP}} \lambda^a) F_P^a \left( 1 - \delta^{1a} \left( 1 + \frac{F_P^8}{F_P^1 \sqrt{2}} \right) \right), \\ &= G_F \frac{m_P^3 m_\ell c_\ell^{\mathcal{P}}}{m_P^2 - M_P^2} \sum_a \text{Tr}(\mathcal{P}^{\text{NP}} \lambda^a) F_P^a \left( 1 - \delta^{1a} \left( 1 + \frac{F_P^8}{F_P^1 \sqrt{2}} \right) \right). \end{aligned} \quad (7.5)$$

<sup>20</sup>At LO,  $\eta_{8(1)} = \eta(\eta') \cos \theta_P \pm \eta'(\eta) \sin \theta_P$ ,  $\theta_P = -19.6^\circ$  [78] and the decay constants read  $F_\eta^8 = F_0 \cos \theta_P$ ,  $F_{\eta'}^8 = F_0 \sin \theta_P$ ,  $F_\eta^1 = F_0 \sin \theta_P$ ,  $F_{\eta'}^1 = F_0 \cos \theta_P$ . In addition,  $\sin(2\theta_P) = \frac{2M_{81}^2}{M_{\eta'}^2 - M_\eta^2}$ , see Ref. [78].

This induces an additional contribution to the  $\mathcal{A}(q^2)$  loop amplitude in Eq. (2.2),

$$\mathcal{A}(q^2) \rightarrow \mathcal{A}(q^2) + \frac{\sqrt{2}G_F m_P^2 c_\ell^{\text{NP}}}{4\alpha^2 F_{P\gamma\gamma}(m_P^2 - M_P^2)} \sum_a \text{Tr}(\mathcal{P}^{\text{NP}} \lambda^a) F_P^a \left( 1 - \delta^{1a} \left( 1 + \frac{F_P^8}{F_P^1 \sqrt{2}} \right) \right).$$

We note that the approximation taken for calculating the  $\langle 0 | \omega | P \rangle$  matrix element has been used with great success in  $J/\Psi \rightarrow \gamma \eta(\eta')$  decays [79]. Actually, at LO in  $\chi\text{PT}$ <sup>21</sup>, the difference between Eq. (7.3) and Eq. (7.4) is of 8%(1%) for the  $\eta(\eta')$ , precise enough for our study. The nice feature from this approach is to provide an appropriate description for the singlet component valid at any order in the chiral expansion and based on known parameters —as long as the approximation employed holds to the required precision.

In the flavor basis, neglecting the  $u$  and  $d$  quark masses, only the strange part contributes. Using an analogous procedure, we find

$$\mathcal{A}(q^2) \rightarrow \mathcal{A}(q^2) + \frac{\sqrt{2}G_F m_P^2 c_\ell^{\text{NP}}}{4\alpha^2 F_{P\gamma\gamma}(m_P^2 - M_P^2)} \sum_a \text{Tr}(\mathcal{P}^{\text{NP}} \lambda^a) F_P^a \left( 1 - \frac{\delta^{as} F_P^q}{\sqrt{2} F_P^s} \right) (1 - \delta^{aq}).$$

All in all, both contributions may be summarized to yield an additional term modifying Eq. (2.2) as

$$\mathcal{A}(q^2) \rightarrow \mathcal{A}(q^2) + \frac{\sqrt{2}G_F F_\pi}{4\alpha_{em}^2 F_{P\gamma\gamma}} (\lambda_P^A + \lambda_P^{\mathcal{P}}), \quad (7.6)$$

where  $G_F$  is the Fermi coupling constant, and  $F_\pi \simeq 92$  MeV is the pion decay constant. The  $\lambda$ -terms depend on the pseudoscalar meson structure, which for the  $\eta$  and  $\eta'$  involve the mixing parameters. In the flavor-mixing scheme, they read<sup>22</sup>

$$\lambda_P^A = c_\ell^A \left[ \frac{F_P^3}{F_\pi} (c_u^A - c_d^A) + \frac{F_P^q}{F_\pi} (c_u^A + c_d^A) + \frac{F_P^s}{F_\pi} \sqrt{2} c_s^A \right], \quad (7.7)$$

$$\lambda_P^{\mathcal{P}} = \frac{c_\ell^{\mathcal{P}}}{1 - \frac{M_P^2}{m_P^2}} \left[ \frac{F_P^3}{F_\pi} (c_u^{\mathcal{P}} - c_d^{\mathcal{P}}) + \frac{F_P^q}{F_\pi} (-c_s^{\mathcal{P}}) + \frac{F_P^s}{F_\pi} \sqrt{2} c_s^{\mathcal{P}} \right]. \quad (7.8)$$

Taking the result from the mixing parameters in Ref. [55] to numerically calculate Eqs. (7.7, 7.8), Eq. (7.6) yields

$$\begin{aligned} \mathcal{A}(m_{\pi^0}^2) &+ 0.026 (c_\ell^A (c_u^A - c_d^A) + c_\ell^{\mathcal{P}} (c_u^{\mathcal{P}} - c_d^{\mathcal{P}}) (1 - M_P^2/m_P^2)^{-1}), \\ \mathcal{A}(m_\eta^2) &+ 0.026 (0.84 c_\ell^A (c_u^A + c_d^A) - 1.27 c_\ell^A c_s^A - 2.11 c_\ell^{\mathcal{P}} c_s^{\mathcal{P}} (1 - M_P^2/m_P^2)^{-1}), \\ \mathcal{A}(m_{\eta'}^2) &+ 0.021 (0.72 c_\ell^A (c_u^A + c_d^A) + 1.61 c_\ell^A c_s^A + 0.89 c_\ell^{\mathcal{P}} c_s^{\mathcal{P}} (1 - M_P^2/m_P^2)^{-1}). \end{aligned}$$

To discuss the sensitivity of each particular channel to NP, it is convenient to cast a very approximate result for  $\mathcal{A}(m_P^2)$ , namely

$$\mathcal{A}(m_P^2) \simeq i\pi \left[ \ln \left( \frac{m_\ell}{m_P} \right) \right] + \left[ \ln^2 \left( \frac{m_\ell}{m_P} \right) - 3 \ln \left( \frac{\Lambda}{m_\ell} \right) + \delta_{\text{NP}} \right], \quad (7.9)$$

<sup>21</sup>At this order, the  $\eta$  and  $\eta'$  masses are [78]  $M_\eta^2 = 0.244 \text{ GeV}^2$ ,  $M_{\eta'}^2 = 0.917 \text{ GeV}^2$  and  $M_0^2 = 0.673 \text{ GeV}^2$ .

<sup>22</sup>By definition,  $F_{\pi^0}^8 = F_{\pi^0}^1 \equiv 0$  and  $F_{\pi^0}^3 \equiv F_\pi$ . From Ref. [56],  $F_{\eta(\eta')}^q = 0.84(0.72)F_\pi$ ,  $F_{\eta(\eta')}^s = -0.90(1.14)F_\pi$  and  $F_{\eta(\eta')}^3 \equiv 0$ . See details in [56] for the errors on these parameters.



where  $\Lambda$  is an effective hadronic scale characterizing the TFF and  $\delta_{\text{NP}}$  is the NP contribution in Eq. (7.6). From Eq. (7.9), we see that, as the lepton mass gets lighter, the amplitude will be dominated by the  $\ln(m_\ell/m_P)$  terms, which become large and make the NP contribution harder to see. Indeed, for  $\ell = e$ , the relative NP contribution to the BR is approximately given by  $2\delta_{\text{NP}}(\ln^2(\frac{m_e}{m_P}) + \pi^2)^{-1}$ . If we are aiming to find contributions from NP, it is therefore much easier to look for the  $\ell = \mu$  channel as the NP part is insensitive to  $m_\ell$  (see Eq. (7.6)).

With respect to  $m_P$ , from the logarithmic scaling, we infer that there is no big difference in the SM in choosing either  $\pi^0$ ,  $\eta$ , or  $\eta'$  as their masses are of same order. Furthermore, the NP axial contribution does not see the  $m_P$ , meaning that is equally likely to appear in any case. This contrasts with the pseudoscalar NP contribution, which strongly depends on  $m_P$  (cf. Eq. (7.8)) and gets bigger as  $m_P$  and  $M_P$  (the mass of the new pseudoscalar particle) approach each other. Still, this is *a priori* irrelevant unless there is a well-motivated NP scale which is close to either the  $\pi^0$ ,  $\eta$ , or  $\eta'$  masses.

From this discussion, we conclude that  $\eta(\eta') \rightarrow \mu^+\mu^-$  decays are the best candidates to look for NP effects (as the  $\pi^0$  cannot decay into muons). For illustrating the statements above, we give the approximate NP contribution to the branching ratio for each particular process,

$$\begin{aligned} BR(\pi^0 \rightarrow e^+e^-) & \left( 1 + 0.001 \left[ c_\ell^A(c_u^A - c_d^A) + c_\ell^P \frac{c_u^P - c_d^P}{1 - M_P^2/m_P^2} \right] \right), \\ BR(\eta \rightarrow_{e^+e^-}^{\mu^+\mu^-}) & \left( 1 + \left( \begin{smallmatrix} -0.002 \\ +0.001 \end{smallmatrix} \right) \left[ 0.84c_\ell^A(c_u^A + c_d^A) - 1.27c_\ell^A c_s^A - \frac{2.11c_\ell^P c_s^P}{1 - M_P^2/m_P^2} \right] \right), \\ BR(\eta' \rightarrow_{e^+e^-}^{\mu^+\mu^-}) & \left( 1 + \left( \begin{smallmatrix} +0.003 \\ +0.001 \end{smallmatrix} \right) \left[ 0.72c_\ell^A(c_u^A + c_d^A) + 1.61c_\ell^A c_s^A + \frac{0.89c_\ell^P c_s^P}{1 - M_P^2/m_P^2} \right] \right). \end{aligned}$$

We see that, as stated above, the  $\ell = e$  channel has the same sensitivity for every pseudoscalar. For  $\ell = \mu$  we find that for  $\eta(\eta')$  is two(three) times more sensitive as the  $\ell = e$  channel. These numbers imply, together with the experimental precision reached for the  $\pi^0(\eta)$  decay (we do not consider the central value, but the obtained precision), bounds for the  $c^A$  parameters of the order  $c^A \sim 7(8)$ . As an example, for the  $Z$  boson ( $c_\ell^A = c_{d,s}^A = -c_u^A \equiv 1$ ), the  $c_f^A$  combination is  $-2(-1.27)(1.61)$  for  $\pi^0(\eta)(\eta')$ .

Interesting enough, a typical  $Z$ -like contribution has opposite sign for  $\pi^0 \rightarrow e^+e^-$  than for  $\eta \rightarrow \mu^+\mu^-$ , contrary to experimental implications. This would suggest either different couplings (necessarily  $SU(2)$  breaking), or lepton flavor violating (LFV) models, which would couple different to distinct generation of quarks, leptons, or both. Moreover, in order to avoid  $(g-2)_\mu$  problems, we would need either some balance from an additional vector-like contribution<sup>23</sup> or, again, LFV models in which the coupling to the muon is suppressed.

For a pseudoscalar contribution, as in Ref. [5], the effective couplings may become even larger as the new particle mass approaches the  $\pi^0, \eta, \eta'$  masses, meaning that would be visible for one of the pseudoscalars alone. Finally, we comment on the existing correlations

---

<sup>23</sup>The dominant Schwinger-like contribution for a vector(scalar)-like coupling has positive sign whereas the axial(pseudoscalar) one has opposite sign, providing a fine tuning cancelation.

given the pseudoscalar structure. We see for instance that  $\pi^0 \rightarrow e^+e^-$  and  $\eta \rightarrow \mu^+\mu^-$  are, in general, anti-correlated unless there is a pseudoscalar particle  $\mathcal{P}$  with  $m_{\pi^0} < m_{\mathcal{P}} < m_\eta$  (or a different structure for distinct generations). Again,  $(g-2)_\mu$  would play an important constraint for the pseudoscalar case as well.

To conclude, there is still the chance to look for NP contributions, specially in the  $\ell = \mu$  channel, and a variety of phenomenology is possible depending on which kind of interaction is chosen. Still, our study suggests to go beyond simple scenarios. This seems nevertheless the standard in high energy physics nowadays, and scenarios of this kind have been and are still studied at present. In this discussion, we have omitted a detailed discussion of available physical constraints for these scenarios. This constitutes a field of study by itself [5, 80–83].

## 8 Conclusions

In this work, we discussed a novel approach for evaluating the Standard Model prediction for  $\eta(\eta') \rightarrow \bar{\ell}\ell$  decays, which are mainly driven by the non-perturbative regime of QCD. This was made possible using the machinery of Canterbury approximants (CA), an extension of Padé approximants to the double virtual case. This approach is data driven, systematic and allows for the correct implementation of the low- and high energy QCD requirements, which are key points in the calculation. In addition, our method implements a systematic error and the results come from a full numerical evaluation for the loop integral, which were not included in most of the previous approaches. From our experience in [11], we expect that higher elements in the approximants sequence yield a value very similar to the OPE  $a_{P,1,1} = 2b_P^2$  choice which also encodes the appropriate high-energy behavior, which means only the correct  $Q^{-2}$  dependence but not its exact coefficient. We quote these as our final number, incorporating the difference with respect to the factorization  $a_{P,1,1} = b_P^2$  choice as an asymmetric systematic error in Table 7. Given the current experimental errors, the achieved precision is accurate enough and does not require the computation of higher elements in the CA sequence, which poses a major complication with respect to the  $\pi^0$  case. We postpone such a calculation for the near future in case new experiments prove to require a higher precision.

Process	BR
$\eta \rightarrow e^+e^-$	$5.31^{(+14)}_{(-4)} \times 10^{-9}$
$\eta \rightarrow \mu^+\mu^-$	$4.72^{(+5)}_{(-21)} \times 10^{-6}$
$\eta' \rightarrow e^+e^-$	$1.82^{(+18)}_{(-18)} \times 10^{-10}$
$\eta' \rightarrow \mu^+\mu^-$	$1.36^{(+29)}_{(-26)} \times 10^{-7}$

**Table 7.** Our final results for the BRs. We take as the central value the OPE result and add as an additional (asymmetric) systematic error the difference with respect to the factorization choice.

In this work, we have found a remarkable feature of our approach, namely, that by abandoning the old conceptions on rational approaches —where the poles are associated to physical resonances and fixed in advance— in favor of the Padé Theory point of view — where poles are effective parameters accounting for the analytic structure of the underlying

function— it is possible to effectively reproduce non-trivial effects encoded in the TFF, such as branch cuts, which is the reason for which we can perform a calculation for the  $\eta(\eta') \rightarrow \bar{\ell}\ell$  decays, beyond the possibilities of resonance approaches.

We have found that, similar to the  $\pi^0 \rightarrow e^+e^-$  case, there is an anomaly for the  $\eta \rightarrow \mu^+\mu^-$  decay, which would make a new measurement very interesting.

In addition, we have discussed the difficulties in achieving a precise description for these processes from  $\chi$ PT, though it would be *a priori* an adequate tool for this purpose; in this respect, we have parametrized the expected higher-order corrections for further studies, which could be useful for LFV studies in  $K_L \rightarrow \bar{\ell}\ell$ . As an outcome, we have parameterized the impact of the  $\pi^0$ -exchange contribution to the  $2S$  hyperfine-splitting in the muonic hydrogen based on our data-driven results.

Finally, we have discussed the implications of New Physics scenarios, which should effectively be of either axial or pseudoscalar nature. Our results suggest that such contributions should arise either from  $SU(2)$  isospin-breaking couplings or lepton flavor violating scenarios.

## A Generalities of Canterbury approximants

In this appendix, we define and illustrate the performance of Canterbury approximants specializing to the case of symmetric functions  $f(x, y) = f(y, x)$ . Given the original function formal expansion

$$f(x, y) = \sum_{\alpha, \beta} c_{\alpha, \beta} x^\alpha y^\beta, \quad (c_{\alpha, \beta} = c_{\beta, \alpha}), \quad (\text{A.1})$$

the Canterbury approximants  $C_M^N(x, y)$  are rational functions of the bivariate degree  $N$  and degree  $M$  polynomials,  $R_N(x, y)$  and  $Q_M(x, y)$ <sup>24</sup>, respectively,

$$C_M^N(x, y) = \frac{R_N(x, y)}{Q_M(x, y)} = \frac{\sum_{k, l=0}^N a_{k, l} x^k y^l}{\sum_{k, l=0}^M b_{i, j} x^i y^j}, \quad (\text{A.2})$$

where the coefficients ( $i \geq j$ )  $a_{i, j} \in \mathcal{N}$  and  $b_{i, j} \in \mathcal{D}$  are determined by the accuracy-through-order conditions ( $b_{0,0} = 1$  as a part of the definition)

$$\sum_{i=0}^{\alpha} \sum_{j=0}^{\beta} b_{i, j} c_{\alpha-i, \beta-j} = a_{\alpha, \beta} \quad \text{for } (\alpha, \beta) \in \mathcal{N}, \quad (\text{A.3})$$

$$\sum_{i=0}^{\min(\alpha, M)} \sum_{j=0}^{\min(\beta, M)} b_{i, j} c_{\alpha-i, \beta-j} = 0 \quad \text{for } (\alpha, \beta) \in \mathcal{E}, (\alpha, \beta) \notin \mathcal{N}, \quad (\text{A.4})$$

where  $\dim(\mathcal{E}) = \dim(\mathcal{N}) + \dim(\mathcal{D}) - 1$ .

---

<sup>24</sup>Note that the polynomials are constructed as to have the maximum power in each variable rather than a total maximum power in  $x^i y^j$  with  $i + j \leq N(M)$  [46]. In other words,  $R_N(x, y)$  contains a term  $x^N y^N$  and  $(N+1)(N+2)/2$  terms in total.

### A.1 Performance of Canterbury approximants with the help of toy models

To illustrate the performance of Canterbury approximants, we take the example of two well-motivated models, the Regge model from Ref. [68] and a logarithmic model which naturally arises for flat distribution amplitudes [84]. These models have been employed in Ref. [68] and Refs. [51, 84] to study the TFF SL data and are selected here given their sophisticated analytic structure. The first one reads

$$F_{P\gamma^*\gamma^*}^{\text{Regge}}(Q_1^2, Q_2^2) = \frac{aF_{P\gamma\gamma} \left[ \psi^{(0)}\left(\frac{M^2+Q_1^2}{a}\right) - \psi^{(0)}\left(\frac{M^2+Q_2^2}{a}\right) \right]}{Q_1^2 - Q_2^2 \psi^{(1)}\left(\frac{M^2}{a}\right)}, \quad (\text{A.5})$$

where  $\psi^n(z) = \partial_z^{n+1} \ln \Gamma(z)$  is the polygamma function, and  $M$  and  $a$  are the parameters for the Regge trajectory  $M_n^2 = M^2 + na$ . For our case of study, we take  $M = 0.8$  GeV and  $a = 1.3$  GeV<sup>2</sup> [85]. The lowest  $C_1^0(Q_1^2, Q_2^2)$  approximant is given, for this model, as

$$C_1^0(Q_1^2, Q_2^2) = \frac{F_{P\gamma\gamma}}{1 - \frac{(Q_1^2+Q_2^2)\psi^{(2)}}{2a\psi^{(1)}} - \frac{Q_1^2Q_2^2}{2a^2} \left( \frac{\psi^{(3)}}{3\psi^{(1)}} - \left(\frac{\psi^{(2)}}{\psi^{(1)}}\right)^2 \right)}, \quad (\text{A.6})$$

where  $\psi^{(n)} \equiv \psi^{(n)}(M^2/a)$ . The second, logarithmic model, reads

$$F_{P\gamma^*\gamma^*}^{\text{log}}(Q_1^2, Q_2^2) = \frac{F_{P\gamma\gamma}M^2}{Q_1^2 - Q_2^2} \ln \left( \frac{1 + Q_1^2/M^2}{1 + Q_2^2/M^2} \right). \quad (\text{A.7})$$

For our study, we take  $M^2 = 0.6$  GeV<sup>2</sup> in analogy to the single-virtual case [51, 84]. Its  $C_1^0(Q_1^2, Q_2^2)$  approximant is then given as

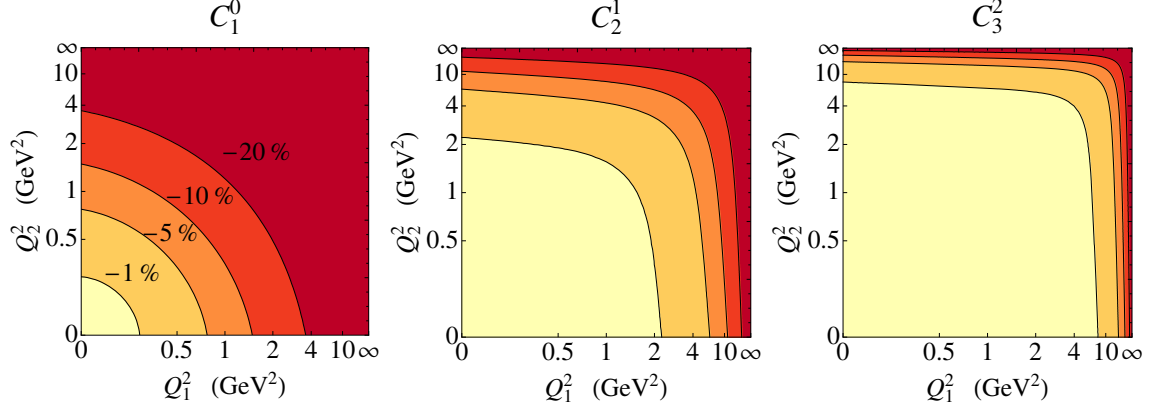
$$C_1^0(Q_1^2, Q_2^2) = \frac{F_{P\gamma\gamma}}{1 + \frac{Q_1^2+Q_2^2}{2M^2} + \frac{Q_1^2Q_2^2}{6M^4}}. \quad (\text{A.8})$$

To show the convergence, we use the  $C_{N+1}^N$  sequence. The reason for this particular choice is not a mathematical but a physical one, namely, that the TFF should vanish as  $Q_{1,2}^2 \rightarrow \infty$ . The results are shown in terms of the relative error  $C_{N+1}^N(Q_1^2, Q_2^2)/F_{P\gamma^*\gamma^*}^{\text{Regge,log}}(Q_1^2, Q_2^2) - 1$  in Figs. 6 and 7 for the Regge and logarithmic model, respectively. The observed convergence pattern is excellent and lies within the expectations of PAs results.

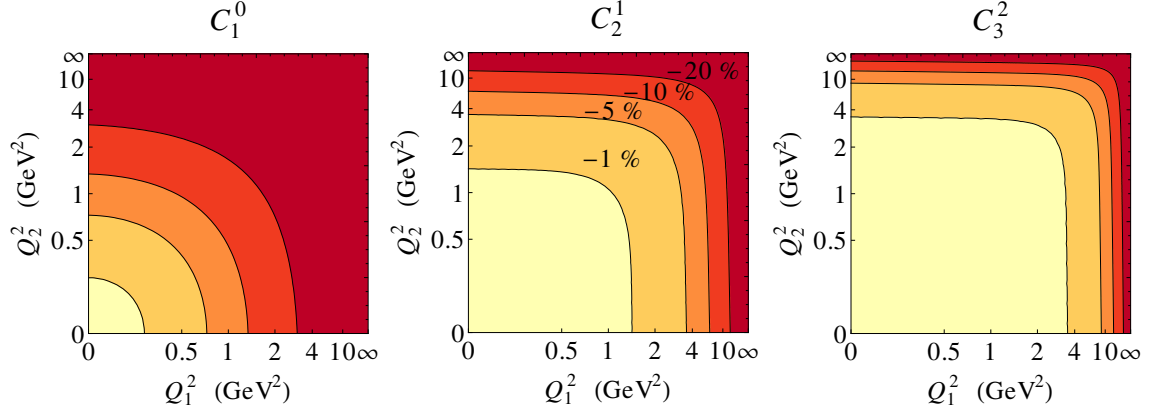
We warn however that fixing the poles in advance to lie at the physical resonances (which holds strictly in the large- $N_c$  limit alone) results in a slower convergence pattern in analogy to PAs [44]. Consequently, we do not advice this practice if only a small set of derivatives are known.

Finally, similar to Padé approximants, the high-energy behavior may be constrained as well. Both of our models fulfill the OPE condition, meaning that, for  $Q_1^2 = Q_2^2 \equiv Q^2 \rightarrow \infty$ , the TFF falls as  $Q^{-2}$ . As an example, for our previous  $C_1^0(Q_1^2, Q_2^2)$  case, this reduces to remove the  $Q_1^2Q_2^2$  term in the denominator. The performance is greatly improved then for  $Q_1^2 = Q_2^2$  up to very large values, at least, for those approximants beyond  $C_1^0(Q_1^2, Q_2^2)$ , which is illustrated in Fig. 8.

In our case of study, we could only reconstruct the  $C_1^0(Q_1^2, Q_2^2)$  approximant given the scarce information on the double-virtual regime. Particularly, we suggested in Ref. [11] that,



**Figure 6.** Convergence of the  $C_{N+1}^N(Q_1^2, Q_2^2)$  sequence to the Regge model for the lowest elements. The first, second, third, and fourth contours, from light to dark red, stand for the relative  $-1$ ,  $-5$ ,  $-10$  and  $-20\%$  deviations. Both axis have been scaled as  $Q^2/(1+Q^2)$ .

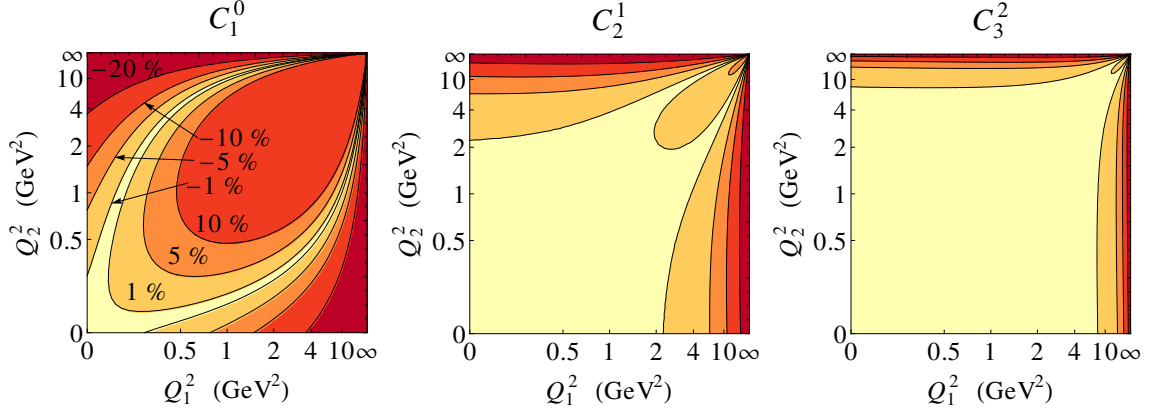


**Figure 7.** Analogous results to those in Fig. 6 but for the logarithmic model.

based on low- and high-energy constraints, the only involved double virtual parameter in this approximant,  $a_{P;1,1}$ , should lie within the  $(b_P^2, 2b_P^2)$  range, which would provide then a conservative estimate for the  $\text{BR}(P \rightarrow \ell\ell)$  calculation. In order to illustrate this statement, we display in Table 8 the range implied for these limiting cases (Fact and OPE columns) for our two given models together with the exact result of the models. When calculating the models results, it is convenient to use alternative expressions equivalent to Eq. (A.5) and Eq. (A.7). Particularly, for the Regge model we use

$$F_{P\gamma^*\gamma^*}^{\text{Regge}}(Q_1^2, Q_2^2) = \frac{F_{P\gamma\gamma}}{\psi^{(1)}(M^2/a)} \sum_{m=0}^{\infty} \frac{a^2}{(Q_1^2 + (M^2 + na))(Q_2^2 + (M^2 + na))}, \quad (\text{A.9})$$

which allows to express the result in terms of a summation of individual terms which can be calculated analog to the factorization case. We find a nice convergence when summing up to  $10^4$  terms after comparison to the  $10^5$  terms summation. For the logarithmic model,



**Figure 8.** Convergence of the  $C_{N+1}^N(Q_1^2, Q_2^2)$  sequence to the Regge model for the first elements as in Fig. 6 but now with the appropriate high-energy behavior imposed. The first, second, third, and fourth outer(inner) contours, from light to dark red, stand for the relative  $\mp 1, \mp 5, \mp 10$  and  $\mp 20\%$  deviations. Both axis have been scaled as  $Q^2/(1+Q^2)$ .

BR( $P \rightarrow \bar{\ell}\ell$ )	Regge				Log		
	Fact	OPE	OPEc	<b>Exact</b>	Fact	OPE	<b>Exact</b>
$\pi^0 \rightarrow e^+e^- \times 10^8$	6.218	6.080	6.266	<b>6.138</b>	5.996	5.869	<b>5.869</b>
$\eta \rightarrow e^+e^- \times 10^9$	4.950	5.064	5.470	<b>5.012</b>	4.614	4.717	<b>4.626</b>
$\eta \rightarrow \mu^+\mu^- \times 10^6$	4.844	5.151	4.688	<b>4.992</b>	5.461	5.889	<b>5.859</b>
$\eta' \rightarrow e^+e^- \times 10^{10}$	1.825	1.781	1.867	<b>1.754</b>	1.469	1.437	<b>1.472</b>
$\eta' \rightarrow \mu^+\mu^- \times 10^7$	1.518	1.407	0.922	<b>1.266</b>	1.419	1.405	<b>1.319</b>

**Table 8.** Our Regge and logarithmic model results for the BR( $P \rightarrow \bar{\ell}\ell$ ) (Exact column) together with their  $C_1^0(Q_1^2, Q_2^2)$  approximant results. The Fact (OPE) column stands for our chosen values in the main text  $a_{P;1,1} = b_P^2(2b_P^2)$ , whereas the OPEc is a choice in which the OPE coefficient is imposed in the  $C_1^0(Q_1^2, Q_2^2)$  reconstruction. See details in the text.

we employ the integral representation

$$F_{P\gamma^*\gamma^*}^{\log}(Q_1^2, Q_2^2) = F_{P\gamma\gamma} M^2 \int_0^1 dx \frac{1}{xQ_1^2 + (1-x)Q_2^2 + M^2}. \quad (\text{A.10})$$

The loop integration can be performed using the same technique as in our OPE case; the resulting value is then integrated over  $x$ . From the results in Table 8, the reader can verify that, for all the cases except the  $\eta'$  ones, the model value lies within the suggested  $a_{P;1,1}$  band. The differences in the  $\eta'$  case are due to the —thoroughly-discussed— appearance of hadronic-induced imaginary parts and are within the error  $\sim 20\%$  that was estimated in Table 3. Summarizing, our error estimations applied to our final results in Table 5 are excellent for the models here investigated.

In addition, the reader may have realized that our  $C_1^0(Q_1^2, Q_2^2)$  description can reproduce the BL and OPE  $Q^2$ -behaviors but not their associated coefficients at once. The latter would be possible with a higher approximant [11], which is however difficult to reconstruct due to our ignorance on double-virtual parameters. Still, given the weight of the

low energies in this process, we argued that this is of no relevance, since the low-energy parameters play the main role. Our particular models chosen here, Eq. (A.5) and Eq. (A.7), do not fulfill the BL  $Q^2$ -behavior (they behave for large  $Q^2$  as  $Q^{-2} \ln Q^2$ ) but they do fulfill the OPE one. As a result, we can illustrate what would have been obtained if our  $C_1^0(Q_1^2, Q_2^2)$  approximant fulfilling the OPE  $Q^2$ -behavior would have been constrained to reproduce the OPE coefficient as well —this is, we trade the slope parameter,  $b_P/m_P^2$ , for the OPE coefficient. Unfortunately, for the logarithmic model this exercise is trivial, since  $F_{P\gamma^*\gamma^*}(Q^2, Q^2) = M^2(M^2 + Q^2)^{-1}$  and a single parameter can account for both, the low and high-energy behaviors. This is not the case for the Regge model, which result when constraining the OPE coefficient is given in the “OPEc” column in Table 8. From this exercise we conclude that, given the weight of the low energies, the most efficient strategy is to rely on the low-energy expansion and the correct high-energy behavior without imposing the high-energy coefficients. The OPEc always represents the worse scenario. This strategy is also supported from a mathematical point of view: whereas the Taylor expansion represents a convergent series with a finite radius of convergence, the high-energy expansion represents only an asymptotic one, for which the convergence of PAs is much slower [44, 45]. Let us finally remark that none of the single entries collected in Table 8 guide towards the exact result. It is only the range given by the two bounds (Fact  $\div$  OPE) which tells where the result is.

## Acknowledgments

Work supported by the Deutsche Forschungsgemeinschaft DFG through the Collaborative Research Center “The Low-Energy Frontier of the Standard Model” (SFB 1044).

## References

- [1] S. Drell, *Direct Decay*  $\pi^0 \rightarrow e^+e^-$ , *Nuovo Cim.* **11** (1959) 693.
- [2] A. Soni, *The Structure of Anomalous Lepton Hadron Interactions and*  $\pi^0 \rightarrow e^-e^+$ , *Phys.Lett.* **B52** (1974) 332.
- [3] P. Herczeg, *On the Decay*  $\pi^0 \rightarrow e^+e^-$ , *Phys.Rev.* **D16** (1977) 712–719.
- [4] Y. Kahn, M. Schmitt and T. M. Tait, *Enhanced rare pion decays from a model of MeV dark matter*, *Phys.Rev.* **D78** (2008) 115002, [0712.0007].
- [5] Q. Chang and Y.-D. Yang, *Rare decay*  $\pi^0 \rightarrow e^+e^-$  *as a sensitive probe of light CP-odd Higgs in NMSSM*, *Phys.Lett.* **B676** (2009) 88–93, [0808.2933].
- [6] PARTICLE DATA GROUP collaboration, K. Olive et al., *Review of Particle Physics*, *Chin.Phys.* **C38** (2014) 090001.
- [7] F. Jegerlehner and A. Nyffeler, *The Muon g-2*, *Phys.Rept.* **477** (2009) 1–110, [0902.3360].
- [8] A. Antognini, F. Nez, K. Schuhmann, F. D. Amaro, F. Biraben et al., *Proton Structure from the Measurement of 2S – 2P Transition Frequencies of Muonic Hydrogen*, *Science* **339** (2013) 417–420.
- [9] A. Dorokhov and M. Ivanov, *On mass corrections to the decays*  $P \rightarrow \ell^+\ell^-$ , *JETP Lett.* **87** (2008) 531–536, [0803.4493].



- [10] A. Dorokhov, M. Ivanov and S. Kovalenko, *Complete structure dependent analysis of the decay  $P \rightarrow \ell^+ \ell^-$* , *Phys.Lett.* **B677** (2009) 145–149, [[0903.4249](#)].
- [11] P. Masjuan and P. Sanchez-Puertas, *Phenomenology of bivariate approximants: the  $\pi^0 \rightarrow e^+ e^-$  case and its impact on the electron and muon g-2*, [1504.07001](#).
- [12] P. Vasko and J. Novotny, *Two-loop QED radiative corrections to the decay  $\pi^0 \rightarrow e^+ e^-$  : The virtual corrections and soft-photon bremsstrahlung*, *JHEP* **1110** (2011) 122, [[1106.5956](#)].
- [13] T. Husek, K. Kampf and J. Novotny, *Rare decay  $\pi^0 \rightarrow e^+ e^-$  : on corrections beyond the leading order*, *Eur.Phys.J.* **C74** (2014) 3010, [[1405.6927](#)].
- [14] T. Husek, K. Kampf and J. Novotny, *Radiative corrections to the Dalitz decay  $\pi^0 \rightarrow e^+ e^- \gamma$  revisited*, *Phys. Rev.* **D92** (2015) 054027, [[1504.06178](#)].
- [15] HADES collaboration, G. Agakishiev et al., *Searching a Dark Photon with HADES*, *Phys.Lett.* **B731** (2014) 265–271, [[1311.0216](#)].
- [16] R. Abegg, A. Baldisseri, A. Boudard, W. Briscoe, B. Fabbro et al., *Measurement of the branching ratio for the decay  $\eta \rightarrow \mu^+ \mu^-$* , *Phys.Rev.* **D50** (1994) 92–103.
- [17] CMD-3 collaboration, R. Akhmetshin et al., *Search for the process  $e^+ e^- \rightarrow \eta'(958)$  with the CMD-3 detector*, *Phys.Lett.* **B740** (2015) 273–277, [[1409.1664](#)].
- [18] M. N. Achasov et al., *Search for the  $\eta' \rightarrow e^+ e^-$  decay with the SND detector*, *Phys. Rev.* **D91** (2015) 092010, [[1504.01245](#)].
- [19] KTeV collaboration, E. Abouzaid et al., *Measurement of the rare decay  $\pi^0 \rightarrow e^+ e^-$* , *Phys.Rev.* **D75** (2007) 012004, [[hep-ex/0610072](#)].
- [20] HADES collaboration, G. Agakishiev et al., *Inclusive dielectron spectra in  $p+p$  collisions at 3.5 GeV*, *Eur. Phys. J.* **A48** (2012) 64, [[1112.3607](#)].
- [21] R. Dzhelyadin, S. Golovkin, V. Kachanov, A. Konstantinov, V. Konstantinov et al., *Study of  $\eta \rightarrow \mu^+ \mu^-$  Decay*, *Phys.Lett.* **B97** (1980) 471.
- [22] P. V. Vorobev et al., *Upper limits of the electron widths of the C-even mesons eta-prime,  $f_0(975)$ ,  $f_2(1270)$ ,  $f_0(1300)$ ,  $a_0(980)$ , and  $a_2(1320)$* , *Sov. J. Nucl. Phys.* **48** (1988) 273–276.
- [23] N. T. Huong, E. Kou and B. Viaud, *Novel approach to measure the leptonic  $\eta^{(\prime)} \rightarrow \mu^+ \mu^-$  decays via charmed meson decays*, [1606.08195](#).
- [24] B.-I. Young, *Rare Decay Modes of the  $\eta$  Meson*, *Phys.Rev.* **161** (1967) 1620–1630.
- [25] C. Quigg and J. Jackson, *Decays of neutral pseudoscalar mesons into lepton pairs*, *UCRL-18487* (1968) .
- [26] K. Babu and E. Ma, *Pseudoscalar Electromagnetic Form-factors: Vector Dominance and Quantum Chromodynamics*, *Phys.Lett.* **B119** (1982) 449.
- [27] L. Ametller, L. Bergstrom, A. Bramon and E. Masso, *The Quark Triangle: Application to Pion and  $\eta$  Decays*, *Nucl.Phys.* **B228** (1983) 301.
- [28] M. Scadron and M. Visinescu, *The Soft Momentum Quark Model Predictions for the Pseudoscalar Meson Decays Into Lepton Pairs*, *Phys.Rev.* **D29** (1984) 911.
- [29] B. Margolis, J. Ng, M. Phipps and H. D. Trottier, *Quark model calculation of  $\eta \rightarrow \ell^+ \ell^-$  to all orders in the bound state relative momentum*, *Phys.Rev.* **D47** (1993) 1942–1950, [[hep-ph/9210259](#)].



- [30] M. J. Savage, M. E. Luke and M. B. Wise, *The Rare decays  $\pi^0 \rightarrow e^+e^-$ ,  $\eta \rightarrow e^+e^-$  and  $\eta \rightarrow \mu^+\mu^-$  in chiral perturbation theory*, *Phys.Lett.* **B291** (1992) 481–483, [[hep-ph/9207233](#)].
- [31] D. Gómez Dumm and A. Pich, *Long distance contributions to the  $K_L \rightarrow \mu^+\mu^-$  decay width*, *Phys.Rev.Lett.* **80** (1998) 4633–4636, [[hep-ph/9801298](#)].
- [32] L. Ametller, A. Bramon and E. Masso, *The  $\pi^0 \rightarrow e^+e^-$  and  $\eta \rightarrow \mu^+\mu^-$  decays revisited*, *Phys.Rev.* **D48** (1993) 3388–3391, [[hep-ph/9302304](#)].
- [33] Z. K. Silagadze, *On the two-photon contributions to  $e^+e^- \rightarrow \eta\gamma$  and  $e^+e^- \rightarrow \eta'\gamma$* , *Phys. Rev.* **D74** (2006) 054003, [[hep-ph/0606284](#)].
- [34] M. Knecht, S. Peris, M. Perrottet and E. de Rafael, *Decay of pseudoscalars into lepton pairs and large  $N(c)$  QCD*, *Phys.Rev.Lett.* **83** (1999) 5230–5233, [[hep-ph/9908283](#)].
- [35] A. E. Dorokhov and M. A. Ivanov, *Rare decay  $\pi^0 \rightarrow e^+e^-$ : Theory confronts KTeV data*, *Phys.Rev.* **D75** (2007) 114007, [[0704.3498](#)].
- [36] A. Dorokhov, *How the recent BABAR data for  $P \rightarrow \gamma\gamma^*$  affect the Standard Model predictions for the rare decays  $P \rightarrow \ell^+\ell^-$* , *JETP Lett.* **91** (2010) 163–169, [[0912.5278](#)].
- [37] S. L. Adler, *Axial vector vertex in spinor electrodynamics*, *Phys.Rev.* **177** (1969) 2426–2438.
- [38] J. Bell and R. Jackiw, *A PCAC puzzle:  $\pi^0 \rightarrow \gamma\gamma$  in the sigma model*, *Nuovo Cim.* **A60** (1969) 47–61.
- [39] G. P. Lepage and S. J. Brodsky, *Exclusive Processes in Perturbative Quantum Chromodynamics*, *Phys.Rev.* **D22** (1980) 2157.
- [40] G.A.Baker and P. Graves-Morris, *Pade Approximants*. Encyclopedia of Mathematics and its Applications, Cambridge Univ. Press, 1996.
- [41] V. A. Novikov, M. A. Shifman, A. I. Vainshtein, M. B. Voloshin and V. I. Zakharov, *Use and Misuse of QCD Sum Rules, Factorization and Related Topics*, *Nucl. Phys.* **B237** (1984) 525.
- [42] L. Bergstrom, E. Masso, L. Ametller and A. Bramon,  *$Q^2$  Duality and Rare Pion Decays*, *Phys.Lett.* **B126** (1983) 117.
- [43] P. Masjuan, E. Ruiz Arriola and W. Broniowski, *Meson dominance of hadron form factors and large- $N_c$  phenomenology*, *Phys. Rev.* **D87** (2013) 014005, [[1210.0760](#)].
- [44] P. Masjuan and S. Peris, *A Rational approach to resonance saturation in large- $N(c)$  QCD*, *JHEP* **05** (2007) 040, [[0704.1247](#)].
- [45] P. Masjuan and S. Peris, *A Rational approximation to  $\langle VV-AA \rangle$  and its  $\mathcal{O}(p^6)$  low-energy constant*, *Phys. Lett.* **B663** (2008) 61–65, [[0801.3558](#)].
- [46] J. S. R. Chisholm, *Rational approximants defined from double power series*, *Mathematics of Computation* **27** (1973) .
- [47] C. Alabiso and P. Butera,  *$N$  - Variable Rational Approximants and Method of Moments*, *J. Math. Phys.* **16** (1975) 840.
- [48] J. S. R. Chisholm and J. McEwan, *Rational approximants defined from power series in  $n$  variables*, *Proc. R. Soc. Lond. A* **336** (1974) 421–452.
- [49] G. Baker, *Essentials of pade approximants, chapter 14, 14.3*. Academic Press, 1975.
- [50] S. Peris, *Large- $N(c)$  QCD and Pade approximant theory*, *Phys. Rev.* **D74** (2006) 054013, [[hep-ph/0603190](#)].

- [51] P. Masjuan,  $\gamma^*\gamma \rightarrow \pi^0$  transition form factor at low-energies from a model-independent approach, *Phys.Rev.* **D86** (2012) 094021, [[1206.2549](#)].
- [52] P. Masjuan and M. Vanderhaeghen, Ballpark prediction for the hadronic light-by-light contribution to the muon  $(g-2)_\mu$ , *J. Phys.* **G42** (2015) 125004, [[1212.0357](#)].
- [53] P. Masjuan, J. J. Sanz-Cillero and J. Virto, Some Remarks on the Pade Unitarization of Low-Energy Amplitudes, *Phys. Lett.* **B668** (2008) 14–19, [[0805.3291](#)].
- [54] P. Masjuan and S. Peris, Pade Theory applied to the vacuum polarization of a heavy quark, *Phys. Lett.* **B686** (2010) 307–312, [[0903.0294](#)].
- [55] R. Escribano, P. Masjuan and P. Sanchez-Puertas, The  $\eta$  transition form factor from space- and time-like experimental data, *Eur. Phys. J.* **C75** (2015) 414, [[1504.07742](#)].
- [56] R. Escribano, S. Gonzalez-Solis, P. Masjuan and P. Sanchez-Puertas, The  $\eta'$  transition form factor from space- and time-like experimental data, [1512.07520](#).
- [57] P. Masjuan, S. Peris and J. Sanz-Cillero, Vector Meson Dominance as a first step in a systematic approximation: The Pion vector form-factor, *Phys.Rev.* **D78** (2008) 074028, [[0807.4893](#)].
- [58] R. Escribano, P. Masjuan and P. Sanchez-Puertas,  $\eta$  and  $\eta'$  transition form factors from rational approximants, *Phys.Rev.* **D89** (2014) 034014, [[1307.2061](#)].
- [59] J. Bijnens, K. Kampf and S. Lanz, Leading logarithms in the anomalous sector of two-flavour QCD, *Nucl.Phys.* **B860** (2012) 245–266, [[1201.2608](#)].
- [60] C. W. Xiao, T. Dato, C. Hanhart, B. Kubis, U. G. Meißner and A. Wirzba, Towards an improved understanding of  $\eta \rightarrow \gamma^*\gamma^*$ , [1509.02194](#).
- [61] L. Landsberg, Electromagnetic Decays of Light Mesons, *Phys.Rept.* **128** (1985) 301–376.
- [62] C. Hanhart, A. Kupsc, U.-G. Meißner, F. Stollenwerk and A. Wirzba, Dispersive analysis for  $\eta \rightarrow \gamma\gamma^*$ , *Eur.Phys.J.* **C73** (2013) 2668, [[1307.5654](#)].
- [63] D. Gómez Dumm, A. Pich and J. Portolés, The Hadronic off-shell width of meson resonances, *Phys.Rev.* **D62** (2000) 054014, [[hep-ph/0003320](#)].
- [64] D. Gómez Dumm and P. Roig, Dispersive representation of the pion vector form factor in  $\tau \rightarrow \pi\pi\nu_\tau$  decays, *Eur.Phys.J.* **C73** (2013) 2528, [[1301.6973](#)].
- [65] P. Masjuan and J. J. Sanz-Cillero, Pade approximants and resonance poles, *Eur. Phys. J.* **C73** (2013) 2594, [[1306.6308](#)].
- [66] G. 't Hooft and M. Veltman, Scalar One Loop Integrals, *Nucl.Phys.* **B153** (1979) 365–401.
- [67] T. Hahn and M. Perez-Victoria, Automatized one loop calculations in four-dimensions and D-dimensions, *Comput.Phys.Commun.* **118** (1999) 153–165, [[hep-ph/9807565](#)].
- [68] E. Ruiz Arriola and W. Broniowski, Pion transition form factor and distribution amplitudes in large- $N(c)$  Regge model, *Phys. Rev.* **D74** (2006) 034008, [[hep-ph/0605318](#)].
- [69] R. Mertig, M. Böhm and A. Denner, FEYN CALC: Computer algebraic calculation of Feynman amplitudes, *Comput. Phys. Commun.* **64** (1991) 345–359.
- [70] L. Arnellos, W. J. Marciano and Z. Parsa, The Decay  $\pi^0 \rightarrow \nu\bar{\nu}\gamma$ , *Nucl.Phys.* **B196** (1982) 365.
- [71] N. T. Huong, E. Kou and B. Moussallam, Single pion contribution to the hyperfine splitting in muonic hydrogen, *Phys. Rev.* **D93** (2016) 114005, [[1511.06255](#)].

- [72] C. Peset and A. Pineda, *The two-photon exchange contribution to muonic hydrogen from chiral perturbation theory*, *Nucl. Phys.* **B887** (2014) 69–111, [[1406.4524](#)].
- [73] M. Ramsey-Musolf and M. B. Wise, *Hadronic light by light contribution to muon  $g-2$  in chiral perturbation theory*, *Phys.Rev.Lett.* **89** (2002) 041601, [[hep-ph/0201297](#)].
- [74] P. Masjuan, *Overview of the hadronic light-by-light contribution to the muon ( $g-2$ )*, *Nucl. Part. Phys. Proc.* **260** (2015) 111–115, [[1411.6397](#)].
- [75] A. Crivellin, G. D’Ambrosio, M. Hoferichter and L. C. Tunstall, *Violation of lepton flavor and lepton flavor universality in rare kaon decays*, *Phys. Rev.* **D93** (2016) 074038, [[1601.00970](#)].
- [76] B. Martin, E. De Rafael and J. Smith, *Neutral kaon decays into lepton pairs*, *Phys.Rev.* **D2** (1970) 179–200.
- [77] R. Escribano, P. Masjuan and J. J. Sanz-Cillero, *Chiral dynamics predictions for  $\eta' \rightarrow \eta\pi\pi$* , *JHEP* **05** (2011) 094, [[1011.5884](#)].
- [78] P. Bickert, P. Masjuan and S. Scherer,  *$\eta$ - $\eta'$  mixing in large- $N_c$  chiral perturbation theory: discussion, phenomenology, and prospects*, *PoS* **CD15** (2016) 056, [[1511.01996](#)].
- [79] R. Escribano and J.-M. Frere, *Study of the  $\eta - \eta'$  system in the two mixing angle scheme*, *JHEP* **06** (2005) 029, [[hep-ph/0501072](#)].
- [80] H. Davoudiasl, H.-S. Lee and W. J. Marciano, *“Dark” Z implications for Parity Violation, Rare Meson Decays, and Higgs Physics*, *Phys. Rev.* **D85** (2012) 115019, [[1203.2947](#)].
- [81] H. Davoudiasl, H.-S. Lee and W. J. Marciano, *Muon Anomaly and Dark Parity Violation*, *Phys. Rev. Lett.* **109** (2012) 031802, [[1205.2709](#)].
- [82] C. E. Carlson and B. C. Rislow, *New Physics and the Proton Radius Problem*, *Phys. Rev.* **D86** (2012) 035013, [[1206.3587](#)].
- [83] S. G. Karshenboim, D. McKeen and M. Pospelov, *Constraints on muon-specific dark forces*, *Phys. Rev.* **D90** (2014) 073004, [[1401.6154](#)].
- [84] A. V. Radyushkin, *Shape of Pion Distribution Amplitude*, *Phys. Rev.* **D80** (2009) 094009, [[0906.0323](#)].
- [85] P. Masjuan, E. Ruiz Arriola and W. Broniowski, *Systematics of radial and angular-momentum Regge trajectories of light non-strange  $q\bar{q}$ -states*, *Phys. Rev.* **D85** (2012) 094006, [[1203.4782](#)].